Al-Ishlah: Jurnal Pendidikan
Vol.16, 3 (September, 2024), pp. 3360-3369
ISSN: 2087-9490 EISSN: 2597-940X, DOI: 10.35445/alishlah.v16i3.5311

Development of Data-Driven Educational Decision Implementation Guidelines Assisted by Formative Tests in Science Learning

Suwito Singgih[1], Agista Sintia Dewi Adila[2], Ismun Nisa Nadhifah[3], Syakira Alfi Kurnia[4], Aza Mazita[5]

[1] Universitas Tidar, Magelang, Indonesia; suwitosinggih@untidar.ac.id
[2] Universitas Tidar, Magelang, Indonesia; agista@untidar.ac.id
[3] Universitas Tidar, Magelang, Indonesia; nisa.nadhifah@untidar.ac.id
[4] Universitas Tidar, Magelang, Indonesia; syakiraalfikurnia25@gmail.com
[5] Universitas Tidar, Magelang, Indonesia; mazitaaza@gmail.com

ARTICLE INFO

Keywords:

DDED;
Formative Tests;
Guides;
Science Learning

Article history:

Received 2024-05-06
Revised 2024-07-18
Accepted 2024-09-24

ABSTRACT

Data-based educational decisions (DDED) are an important part of the learning process, and they are used to improve the quality of learning. This research purpose to develop Data Driven Educational Decision Implementation Guidelines (DDED) supporting formative tests in IPA learning. DDED is an adaptation of Data Drive Decision Making (DDDM) in learning. DDED is based on data drive decision making (DDDM), which is a decision-making process that is widely applied in the world of management. The DDED directs teachers to be adaptive and flexible in implementing various methods, models, teaching materials and learning media based on student access. The development design used is 4D, consisting of defining, designing, developing and deploying. The content of the guidelines has been agreed through expert consensus through the fuzzy delphi technique. The fuzzy Delphi method is used to create a broad and detailed understanding, especially in education. The consensus indicated that formative tests should include relevant questions and quizzes, comprising both multiple-choice and essay formats, aligning with the principles of DDED. The resulting guideline tests the level of validity and user response. Users respond well to the guidelines developed.

Corresponding Author:
Agista Sintia Dewi Adila
Universitas Tidar, Magelang, Indonesia; agista@untidar.ac.id

1. INTRODUCTION

Data Driven Educational Decision (DDED) refers to educational professionals who systematically collect and analyze data both before, at the time and after learning to improve the quality of learning (Downey, C et al., 2013 & Cosner, S 2012). DDED is rooted in data driven decision making (DDDM) which has been widely applied by management professionals in various corporations and organizations to improve performance based on empirical data. However, DDED has not been much applied to the world of education. (Cosner, S. 2012 & Marsh, J.A., et al 2006). While the teacher is the manager in classroom learning. Although the interest and adoption of DDED

by educational professionals is increasing (Lai, M.K et al., 2013 & Prenger, R., et al. 2018), there is not much research done on the effect of the DDED on educators in adapting learning to improve student learning outcomes. (Lai, M.K et al., 2016 & Lai, M.K et al., 2013). Many experts also suggest that it may be difficult for teachers to implement DDED in learning planning because many teachers are unfamiliar and less skilled when interacting with data. (Hoogland, I et al., 2016, Reeves, T.D et al., 2016, Von der Embse, N.P et al., 2017, Prenger, R et al., 2018 & Achen, J et al., 2018). The DDED concept is an inseparable part of teachers because basically teachers must continue to improve the quality of their learning so that they are able to improve students' abilities after the learning process.

There have been many studies that report positive effects of the implementation of formative assessments on students as these assessments help students reflect their knowledge, encourage and train to ask relevant questions, think critically, predict and regulate their own learning processes (Zainudin, Z et. al., 2022, Song, Y et al., 2019, Hondrich, A.L et al., 2016). This positive impact also exists on teachers to identify the difficulties experienced by students so that they can always adjust learning based on the level of knowledge and needs of students. (Cisterna, D et al., 2018, Nadhifah, I.N et al., 2023). Assessments performed with the aim of improving learning must be different from those used to simply monitor progress and determine learning success (Nadhifah, I.N, 2012). Scientific learning assessments in Indonesia still tend to be merely summary assessments that measure more quantitative understanding than the underlying qualitative concepts. (Monteiro, V et al., 2021, Nadhifah, I.N, 2012). In fact, to make students affectionate and engage in science, the scaffolding of their knowledge must be based on a strong understanding of qualitative concepts (Nadhifah, I.N, 2012).

A formative test is used to measure the student's ability to master a defined competence in determining the DDED decision (Du, Y. 2022). In order for the student to gain a complete understanding then, the teacher must always adjust the strategy and purpose of his study to the conditions of the student's dynamic understanding. Then the teacher should be able to identify such things from the right data and analysis at the right time. This is where the knowledge and skills of teachers in using strategies and formative assessment tools that fit their class conditions are essential to obtain empirical data. This data is analyzed and referenced as a basis for teacher decision making in adjusting their learning strategy (Data Driven Educational Decision (DDED)) so that the learning process becomes dynamic to accommodate the needs of students (Datnow, A et al., 2013). Therefore, it is appropriate for DDED to be a model used by teachers to improve their learning process.

The purpose of this research is to develop guidelines for the application of DDED supported by formative assessments in IPA learning. The introduction identifies gaps in current research and practice but could benefit from a more explicit statement highlighting the specific gap this study aims to address. For instance, "Despite increasing interest in DDED, there is a lack of practical guidelines for teachers to effectively implement data-driven decisions in classroom settings." (Nadhifah, I.N et al., 2023, Pham, T et al., 2021 & McCallum, S et al., 2021). From this research, it is clear that in designing formative assessments, teachers need practical concepts and this can be facilitated through the DDED concept. The urgency of the teacher's ability to apply the correct formative assessment strategies and tools, analyze the data obtained and design the adaptation of learning as a follow-up of the results becomes very important for students to build an integrated IPA concept so as to cultivate a sense of ownership of the knowledge of IPA and motivated in the field of STEM in the future. (Black, P et al., 2009). One of the key factors in realizing this is the development of professional pedagogical competence of teachers, especially in the assessment aspects and the design of learning strategies (Kayan, F.F et al., 2022). The urgent problem that needs to be addressed is the need for practical guidelines with steps that can be followed by teachers easily to prepare teachers to become skilled learning analysts and designers. The guideline is expected to be a practical guide for teachers in collecting learning data, analyzing and designing appropriate follow-up strategies.

2. METHODS

The research was conducted by adopting the 4D development model (Thiagarajan, 1974), which comprises defining, designing, developing, and disseminating. The development of the DDED implementation guidelines, supported by formative tests in science learning, was limited to the development stage.

The first stage is to define, aiming to analyze the needs and identify the necessary information for development. Needs analysis aims to identify facts and issues in implementing DDED. The facts and issues discovered then serve as the basis for establishing the content in the guideline book. The fuzzy Delphi technique is used in this stage (Saido, 2018). The steps of the fuzzy Delphi technique involve: 1) selecting 10 sources consisting of science educators, formative assessment experts, and researchers in science education; 2) conducting interviews with experts related to the facts and issues of implementing DDED in science learning; 3) analyzing interview results for transcription and establishing guideline book content referring to literature review findings; 4) seeking expert assessments on the guideline book content using a questionnaire; 5) analyzing questionnaire results using the fuzzy Delphi technique.

The fuzzy Delphi technique analysis is conducted following these steps:
1) Determining the linguistic scale. The linguistic scale utilizes a Likert scale by incorporating three fuzzy numbers (m1, m2, m3) ranging from 0 to 1, as presented in Table 1.

Table 1. Linguistic Scale with Fuzzy Numbers

Linguistic Scale	Fuzzy Number		
	m_1	m_2	m_3
Strongly agree	0,6	0,8	1
Agree	0,4	0,6	0,8
Somewhat agree	0,2	0,4	0,6
Disagree	0	0,2	0,4
Strongly disagree	0	0	0,2

2) Calculating the average score of fuzzy numbers ($\overline{(FN)}$)
3) Computing the threshold value (d) using equation (1).

$$d = \sqrt{\frac{1}{3}(M_1 - m_1)^2 + (M_2 - m_2)^2 + (M_3 - m_3)^2} \quad (1).$$

4) calculate the percentage value of expert consensus; and
5) compute the alpha-cut value (DV) using equation (2).

$$DV = \frac{1}{3}(m_1 + m_2 + m_3) \quad (2).$$

Expert consensus is obtained if the value of d ≤ 0.2, the percentage is ≥75%, and the alpha-cut (DV) > 0.5 (Murry & Hammons, 1995; Cheng & Lin, 2002; Chu & Hwang, 2008).

After finalizing the agreed-upon content, the next phase was the design stage. In this phase, the content established during the define phase served as the foundation for developing the guidelines. The structure and format of the guideline were systematically determined during this stage. The design phase resulted in a draft of the DDED implementation guideline, which was further refined through formative assessments in the context of learning science.

After the draft of the book was finished, the next phase was developed. The development phase is carried out by testing the validity and user test or user's response. The way to determine the validity of a test is based on the results of responses from 10 experts. Then, the validity score was tested using V Aiken by looking at the number of respondents using a Likert scale of 1 to 5. Data analysis of the validity test questionnaire is performed using equation (3).

$$V = \frac{\sum(r - l_o)}{[n(c-1)]} \quad (3)$$

The result of validity test is considered valid if the value of V ≥ 0.70 (Aiken, 1996). Meanwhile, the way to determine user responses is by distributing questionnaires to 20 respondents who are junior high school science teachers. The analysis of user response questionnaires was conducted using equation (4).

$$score\ (\%) = \frac{total\ score}{maximal\ score} x100\% \qquad (4).$$

The decision on user response is based on the user response assessment categories presented in Table 2 (Arikunto, 2009).

Table 2. User Response Criteria

Score (%)	Criteria
<21%	Very Poor
21% - 40%	Poor
41% - 60%	Fair
61% - 80%	Good
81% - 100%	Excellent

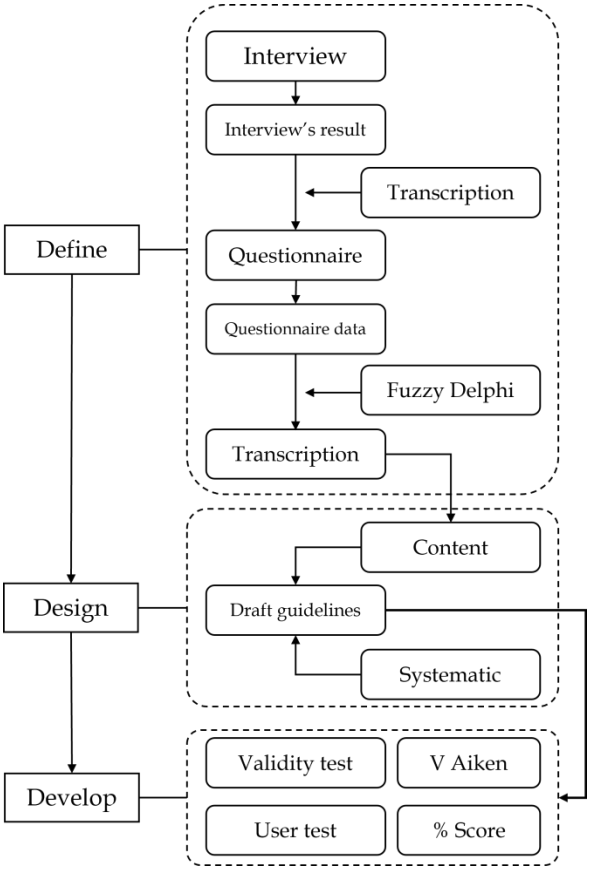

Figure 1. Research Flowchart

3. FINDINGS AND DISCUSSION

3.1 Results of Content Determination Analysis using the Fuzzy Delphi Method

This research aims to develop guidance for implementation of DDED with the help of formative tests in IPA learning. The 4D development model is carried out, with limitation to the dovelop phase.

The Stage Define, is aimed at identifying facts and problems in the field that are later used as the basis for setting the content required in the guidelines. The fuzzy analysis of delphi is carried out to obtain consensus among experts. The results of the analysis and the content agreed in the guidelines are shown in table 3, with DV ≥ 0.5.

Table 3. Guide Content

Findings	Content	Value *d*	Presentation of agreement
Teachers don't understand DDED	Meaning of DDED	0,17	88%
Teachers use intuition in determining the learning access of each student in the learning process Teachers are not used to using formative tests in learning	Formative test	0,19	82%
Teachers conduct formative tests but not as a form of DDED implementation The formative test is commonly used by teachers in the form of quizzes and oral questions.	Integration of formative tests in DDED	0,15	88%
Teachers set learning success based on KKM or KKTP	KKM or KKTP as the basis for DDED decision-making	0,19	82%
Teacher difficulties in assembling learning tools that implement DDED Less learning time A considerable number of students	DDED integration with formative tests in contextual learning	0,05	88%

Based on the consensus results, the guidelines were developed during the design phase, where the systematic structure and format were established. This research addresses the need for the implementation of DDED in learning. Differential learning, as defined by VanTassel-Baska (2012), tailors teaching methods to meet the unique needs of each student and the way they learn. Through DDED, teachers can better identify and accommodate the individual learning processes of their students. The research consensus suggests that DDED can be effectively implemented by incorporating formative assessments consistently within learning activities. Formative tests, as described by Gronlund (1985), monitor student progress and provide ongoing feedback to both students and teachers regarding learning successes and areas for improvement. In the context of DDED, relevant formative assessments include feedback through oral questioning, as well as test formats such as multiple choice, true/false, and short answer questions. Additionally, observation can be used as a formative tool, although it is recommended that an observer, other than the teacher, assist in observing student behaviors and interactions to ensure more objective insights.

The consensus states that, in making DDED decisions, it is possible to review formative test results against the Minimum Mastery Criteria (KKM) or Learning Objective Achievement Criteria (KKTP). KKM or the Minimum Mastery Criteria is the standard of learning achievement adjusted by educational institutions based on the competency criteria that students must possess upon graduation or the graduate competency standards (SKL) in the 2013 Curriculum, as cited by (Hermawati et al., 2018). On the other hand, KKTP or Learning Objective Achievement Criteria are a set of criteria or guidelines describing how well learners have mastered the targeted competencies within the learning objectives in the Merdeka Curriculum (Rosyida et al., 2023). Determining KKM or KKTP can be set by teachers considering the general ability conditions of students (reference). The implementation of DDED directs teachers to be adaptive in implementing various methods or models, materials, and learning media. Teachers are flexibly able to change the direction of learning activities when it is found that students are unable to follow the ongoing learning activities. Through formative tests, teachers periodically monitor and evaluate students' learning achievements.

In making decisions within DDED (Data-Driven Educational Decision-making), the number of students who have met and exceeded the Minimum Competency Level (KKM) or Minimum Completeness Criteria (KKTP) becomes a consideration. KKM is the limit score that students must achieve after studying. Consensus states that a DDED decision can be established when the number of students who have met the KKM or KKTP is between 60-70% of the total number of students in the class. This determination is based on the perspective that learning is successful if the number of students achieving the minimum learning objectives is at least 80% (Gulo, 2022). DDED is a systematic process of collecting and analyzing data oriented toward improving learning activities. The implementation of DDED in the learning process is based on the students' achievement towards the learning objectives. DDED guides teachers to observe when the learning process is no longer effective and when students struggle to achieve the learning objectives. Teachers are required to be able to recognize when these conditions occur, not solely based on intuition as educators but grounded in relevant data. Achieving the KKM or KKTP in formative tests can serve as valid data for teachers to evaluate the success of the learning process.

3.2 Validity Test Results

The Implementation Guideline of DDED Assisted by Formative Tests, Accompanied by Examples of its Application in Contextual Learning. Contextual learning is relevant to science education, where scientific studies are contextual (Gita, 2018). Various methods or models of contextual learning align with the application of DDED. Formative test data collection can be conducted at each phase or syntactic stage of learning. Teachers can continuously monitor students' achievements in participating in learning activities. With evaluation at each stage of the learning activities, teachers can make appropriate improvements, thereby achieving the learning objectives.

The draft guidelines that have been generated are subsequently tested for validity and user response in the development stage. The validity test results indicate that the guidelines are valid, as shown in Table 4.

Tabel 4. Validity Test Results

Indicator	Value V	Criteria
Content Suitability		
1. Relevance to guideline content	0,825	Valid
2. Accuracy of theories	0,750	Valid
3. Up-to-dateness of content	0,800	Valid

Presentation Suitability			
1.	Accuracy in PUEBI	0,875	Valid
2.	Ease of language comprehension	0850	Valid
3.	Communicative language	0,775	Valid
4.	Clear and coherent presentation	0,775	Valid
5.	Attractive presentation	0,750	Valid

The results of the user response test are presented in Table 5.

Table 5. User Response Test Results

Indikator	%	Kriteria
Attractive design	85	Excellent
Consistent selection of illustrations, colors, and fonts	80	Good
Compliance with PUEBI guidelines	86	Excellent
Understandable sentences in guidelines	88	Excellent
Communicative and interactive language	80	Good
Easily understandable content	86	Excellent
No ambiguity	82	Excellent
Easily understandable terminology	85	Excellent
Clear and coherent material presentation	88	Excellent
Easy-to-use guidelines	84	Excellent

It is essential for teachers to implement DDED with the support of formative assessments. However, DDED has not yet been widely adopted by many educators, highlighting the need to raise awareness and promote its application among science teachers to improve the quality of instruction. Despite its proven benefits, DDED has not seen widespread implementation in educational settings (Cosner, 2012; Marsh et al., 2006). As classroom managers, teachers play a crucial role in adopting innovative teaching strategies like DDED. While there is growing interest in DDED among educational professionals (Lai et al., 2013; Prenger et al., 2018), it requires active participation from all levels of the educational system to fully integrate this approach. Enhancing the science learning process through DDED will allow students to gain a deeper understanding and improve their overall academic performance.

4. CONCLUSION

This research aims to develop guidelines for implementing DDED with the assistance of formative tests in teaching science. The 4D development model is employed in this study, focusing

specifically on the development phase. This development is grounded in field findings and facts indicating the challenge of implementing DDED in science teaching. Overall, the difficulties stem from teachers' limited understanding of DDED implementation, hence the necessity for guidelines to serve as a reference for teachers applying DDED, particularly in science education. The guideline's content is the outcome of expert consensus obtained through the fuzzy Delphi technique. The results of validity testing and user responses indicate that the formulated guideline is both valid and practical for use. There is a need for further research on the DDED concept and the dissemination of the DDED concept to teachers in order to improve the quality of their learning.

REFERENCES

Achen J, Harrison C, Cohen-Vogel L. Data Use as Instructional Reform: Exploring Educators' Reports of Classroom Practice. Leadersh Policy Sch [Internet]. 2018 Apr 3;17(2):296–325. Available from: https://www.tandfonline.com/doi/full/10.1080/15700763.2016.1278244

Aiken, L. R. (1996). *Rating scales and checklists: Evaluating behavior, personality, and attitudes. John Wiley & Sons.*

Arikunto, S. (2009). *Dasar-Dasar Evaluasi Pendidikan Edisi (Revisi)*. Jakarta: Bumi Aksara.

Black P, Wiliam D. Developing the theory of formative assessment. Educ Assessment, Eval Account [Internet]. 2009 Feb 23;21(1):5–31. Available from: http://link.springer.com/10.1007/s11092-008-9068-5

Cheng, C. H., & Lin, Y. (2002). Evaluating the best main battle tank using fuzzy decision theory with linguistic criteria evaluation. *European journal of operational research, 142*(1), 174-186. https://doi.org/10.1016/S0377-2217(01)00280-6

Chu, H. C., & Hwang, G. J. (2008). A Delphi-based approach to developing expert systems with the cooperation of multiple experts. *Expert systems with applications, 34*(4), 2826-2840. https://doi.org/10.1016/j.eswa.2007.05.034

Cisterna D, Gotwals AW. Enactment of Ongoing Formative Assessment: Challenges and Opportunities for Professional Development and Practice. J Sci Teacher Educ. 2018 Apr 3;29(3):200–22.

Cosner S. Leading the Ongoing Development of Collaborative Data Practices: Advancing a Schema for Diagnosis and Intervention. Leadersh Policy Sch [Internet]. 2012 Jan;11(1):26– 65. Available from: http://www.tandfonline.com/doi/abs/10.1080/15700763.2011.577926

Datnow A, Park V, Kennedy-Lewis B. Affordances and constraints in the context of teacher collaboration for the purpose of data use. J Educ Adm [Internet]. 2013 May 3;51(3):341–62. Available from: https://www.emerald.com/insight/content/doi/10.1108/09578231311311500/full/html

Downey, Christopher.; Kelly A. Professional attitudes to the use of data in England. In: Data-based Decision Making in Education: Challenges and Opportunities. 2013. p. 69–89.

Du, Y. (2022). Application of the Data-Driven Educational Decision-Making System to Curriculum Optimization of Higher Education. *Wireless Communications and Mobile Computing, 2022*, 1-8.

Gita, S. D., Annisa, M., & Nanna, W. I. (2018). Pengembangan modul IPA materi hubungan makhluk hidup dan lingkungannya berbasis pendekatan kontekstual. *LENSA (Lentera Sains): Jurnal Pendidikan IPA, 8*(1). https://doi.org/10.24929/lensa.v8i1.28

Gronlund, Norman E. dan Robert L. Linn. 1985. Measurement and Evaluation in Teaching. New York: Macmillan Publising Company.

Gulo, A. (2022). Penerapan Model Discovery Learning Terhadap Hasil Belajar Peserta Didik Pada Materi Ekosistem. *Educativo: Jurnal Pendidikan, 1*(1), 307-313. https://doi.org/10.56248/educativo.v1i1.54

Hermawati, H., Nurcahyono, N. A., & Setiani, A. (2018). Proses Pelaksanaan Remedial Teaching Terhadap Ketuntasan Belajar Matematika Peserta Didik. *SOSIOHUMANIORA: Jurnal Ilmiah*

Ilmu Sosial Dan Humaniora, *4*(2). https://doi.org/10.30738/sosio.v4i2.2823

Hondrich AL, Hertel S, Adl-Amini K, Klieme E. Implementing curriculum-embedded formative assessment in primary school science classrooms. Assess Educ Princ Policy Pract. 2016 Jul 2;23(3):353–76.

Hoogland I, Schildkamp K, van der Kleij F, Heitink M, Kippers W, Veldkamp B, et al. Prerequisites for data-based decision making in the classroom: Research evidence and practical illustrations. Teach Teach Educ [Internet]. 2016 Nov;60:377–86. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0742051X16301408

Kayan-Fadlelmula F, Sellami A, Abdelkader N, Umer S. A systematic review of STEM education research in the GCC countries: trends, gaps and barriers. Int J STEM Educ [Internet]. 2022 Jan 3;9(1):2. Available from: https://stemeducationjournal.springeropen.com/articles/10.1186/s40594-021-00319-7

Lai, Mei Kuin; Schildkamp K. In-Service Teacher Professional Learning: Use of Assessment in Data-based Decision-making. In: Handbook of Human and Social Conditions in Assessment. 2016. p. 18.

Lai MK, McNaughton S. Analysis and Discussion of Classroom and Achievement Data to Raise Student Achievement. In: Data-based Decision Making in Education [Internet]. Dordrecht: Springer Netherlands; 2013. p. 23–47. Available from: http://link.springer.com/10.1007/978-94-007-4816-3_3

Lai MK, Schildkamp K. Data-based Decision Making: An Overview. In: Data-based Decision Making in Education [Internet]. Dordrecht: Springer Netherlands; 2013. p. 9–21. Available from: http://link.springer.com/10.1007/978-94-007-4816-3_2

Marsh, Julie A.; Pane, John F.; Hamilton LS. Making Sense of Data-Driven Decision Making in Education: Evidence from Recent RAND Research [Internet]. RAND Corporation; 2006. Available from: https://www.rand.org/pubs/occasional_papers/OP170.html

McCallum S, Milner MM. The effectiveness of formative assessment: student views and staff reflections. Assess Eval High Educ [Internet]. 2021 Jan 2;46(1):1–16. Available from: https://www.tandfonline.com/doi/full/10.1080/02602938.2020.1754761

Monteiro V, Mata L, Santos NN. Assessment Conceptions and Practices: Perspectives of Primary School Teachers and Students. Front Educ [Internet]. 2021 Apr 7;6. Available from: https://www.frontiersin.org/articles/10.3389/feduc.2021.631185/full

Murry Jr, J. W., & Hammons, J. O. (1995). Delphi: A versatile methodology for conducting qualitative research. *The review of higher education*, *18*(4), 423-436. https://doi.org/10.1353/rhe.1995.0008

Nadhifah IN. Identifying Indonesian Students Qualitative Conceptions of Force and Motion. University College London; 2016.

Nadhifah IN. Pengembangan Perangkat Penilaian Afektif dan Karakter pada Pembelajaran Fisika Untuk Sekolah Menengah Atas [Internet]. UIN SUNAN KALIJAGA.; 2012. Available from: https://digilib.uin-suka.ac.id/id/eprint/7249/

Nadhifah, Ismun Nisa; Adila, Agista Sintia Dewi ; Lestari A. Deskripsi Pemahaman, Persepsi dan Kendala Terhadap Penerapan Asesmen Formatif pada Guru IPA di Wonosobo. PENDIPA J Sci Educ [Internet]. 2023;7(1):14–23. Available from: https://ejournal.unib.ac.id/pendipa/article/view/26148

Pham T, Pham LH. Implementing Formative Assessment in Vietnamese Classrooms: Strategies to Navigate Cultural and Structural Obstacles. In 2021. p. 137–50. Available from: https://link.springer.com/10.1007/978-981-16-0057-9_9

Prenger R, Schildkamp K. Data-based decision making for teacher and student learning: a psychological perspective on the role of the teacher. Educ Psychol [Internet]. 2018 Jul 3;38(6):734–52. Available from: https://www.tandfonline.com/doi/full/10.1080/01443410.2018.1426834

Reeves TD, Summers KH, Grove E. Examining the landscape of teacher learning for data use: The

case of Illinois. Boylan M, editor. Cogent Educ [Internet]. 2016 Dec 31;3(1):1211476. Available from: https://www.tandfonline.com/doi/full/10.1080/2331186X.2016.1211476

Saido, G. A. M., Siraj, S., DeWitt, D., & Al-Amedy, O. S. (2018). Development of an instructional model for higher order thinking in science among secondary school students: a fuzzy Delphi approach. *International Journal of Science Education*, *40*(8), 847-866. https://doi.org/10.1080/09500693.2018.1452307

Song Y, Sparks JR. Building a game-enhanced formative assessment to gather evidence about middle school students' argumentation skills. Educ Technol Res Dev [Internet]. 2019 Oct 28;67(5):1175–96. Available from: http://link.springer.com/10.1007/s11423-018-9637-3

Rosyida, N. I., Nugraha, S. W., Agustin, U., Rochman, C., & Nuryantini, A. Y. (2023). Optimalisasi Program Remedial Berdasarkan Instrumen Akreditasi Satuan Pendidikan Dalam Pembelajaran Diferensiasi Kurikulum Merdeka. *Jurnal Evaluasi Pendidikan*, *14*(1), 12-16. https://doi.org/10.21009/jep.v14i1.36763

Thiagarajan, S. (1974). *Instructional development for training teachers of exceptional children: A sourcebook.*

VanTassel-Baska, J. (2012). Analyzing Differentiation in the Classroom. Gifted Child Today, 35(1), 42–48. https://doi.org/10.1177/1076217511427431

von der Embse NP, Schoemann AM, Kilgus SP, Wicoff M, Bowler M. The influence of test-based accountability policies on teacher stress and instructional practices: a moderated mediation model. Educ Psychol [Internet]. 2017 Mar 16;37(3):312–31. Available from: https://www.tandfonline.com/doi/full/10.1080/01443410.2016.1183766

Zainuddin Z, Shujahat M, Haruna H, Chu SKW. The role of gamified e-quizzes on student learning and engagement: An interactive gamification solution for a formative assessment system. Comput Educ [Internet]. 2020 Feb;145:103729. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0360131519302829