



ANALYSIS OF NON-FORMAL EDUCATION TUTOR CAPABILITIES IN EXPLORING ASSESSMENT FOR SCIENCE LEARNING

D. Darmawan¹, D. Yatimah², K. Sasmita³, R. Syah*⁴

^{1,2,3,4}Universitas Negeri Jakarta

DOI: 10.15294/jpii.v9i2.24025

Accepted: April 14th 2020. Approved: June 25th 2020. Published: June 30th 2020

ABSTRACT

This study aims to explore cognitive assessment techniques based on teaching experience. The method in this study uses a descriptive survey with data collection techniques using purposive sampling of tutors who have taught at PKBM for 5 years. The population in this study were PKBM (*Pusat Kegiatan Masyarakat* or Community Learning Center) tutors in Jakarta at Package A, Package B, and Package C, namely 158 tutors with details as follows; Package A 45 tutors, Package B 53 Tutors, and Package C 60 tutors. Data collection was done by using question forms and data analysis used descriptive translation to obtain an average frequency of learning assessment techniques. The statistical analysis used Confirmatory Factor Analysis, so that it can be seen directly and indirectly. The results showed that the dominant assessment technique used in non-formal science tutors was a written test that averaged 32 times in one semester. A group of tutors with 5 to 10 years of teaching experience using a variety of assessment techniques. Descriptively, it can be said that the use of the frequency spectrum of cognitive assessment techniques is coupled with experienced tutors with a tendency in written tests. In this study, the results show that the preparation of assessment in non-formal education there is an influence between the preparation of the assessment of the implementation of the assessment. So it is necessary to improve tutor assessment preparation.

© 2020 Science Education Study Program FMIPA UNNES Semarang

Keywords: cognitive assessment of science; teaching experience; diverse assessment techniques

INTRODUCTION

Assessment is an important part of learning that is used to provide information as a material consideration to determine the success or failure of a learning system. De Vries et al. (2014) explains that assessment is used by educators and schools to determine student learning success. Not only used in formal education, but an assessment is also used in non-formal education such as Community Learning Centers or PKBM. In Non-formal Education, the Tutor as an educator in PKBM is the person responsible for improving student learning performance in the classroom (Sandi et al., 2018). The level of mastery competency tutors is very influential in learning (Copriady, 2014). One of the competencies in question

is the ability of tutors in learning management including assessment management. To be a competent tutor in terms of assessment requires a good academic background and adequate teaching experience. Tutor's background influences learning and assessment in the classroom (Borg, 2018).

Tutors play an integral role in evaluating students, for this reason, their competency and knowledge skills in classroom assessment practice need to be developed (Borg, 2018).

Tutor competence in exploring assessment as a series of characteristics, knowledge, skills, and attitudes of tutors needed to improve effective tutor performance in the educational context (De Vries et al., 2014). Tutors understand the skills of assessment practice in class reflecting their perception of their skills in carrying out classroom assessment practices. Also, Tutors can assess their assessment skills, so that if they can carry

*Correspondence Address
E-mail: rahmat29syah@gmail.com

out assessment practices to the maximum (Offerdahl & Tomanek, 2011). When asked about the assessment of their classroom practices, tutors generally show that they are less proficient in developing evaluation and assessment systems (Koloi-Keaikitse, 2016). As educators, tutors are aware of the implications of the assessment system in the field of teaching and learning in non-formal education (Blazar & Kraft, 2017). Inadequate curriculum training and tutor assessments can influence competency-based assessments needed for sustainable development (Brandt et al., 2019). Therefore, it is important to develop tutor assessment skills towards improving classroom assessment practices to address student needs.

One of the determinants of a tutor's professional attitude is the tutor's experience in assessment because this allows the tutor to inform how various types of reviews are used or misused and what can be done to improve assessment practices in the classroom. Koloi-Keaikitse (2016) supports the need to develop tutor competencies in assessment because if tutors feel ready when they enter the world of education, they are likely to have better assessment skills that may have effects that will ultimately increase their motivation to teach.

A PKBM tutor must be able to master the domain of assessment in the Indonesian curriculum, which is the 2013 curriculum which includes the spiritual domain, social attitudes, knowledge, and skills. More generally it can be categorized into three domains namely cognitive (knowledge), affective (social and spiritual attitudes), and psychomotor (skills). Evaluation standards by educators according to BSNP include general standards, planning standards, implementation standards, processing standards, and reporting of assessment results as well as standards for the use of assessment results (Raharjo, 2012).

The use of assessment techniques is adjusted to the needs that can support teaching programs such as basic competencies to be achieved (Hooper et al, 2014). Careful planning such as making instrument lines is expected to provide accurate information about student competencies that need to be measured, encourage learners to learn more actively to improve their competitiveness, motivate teaching educators to improve student competencies, improve institutional performance and improve quality education (Darling-Hammond et al., 2020). In other words, assessment can be used to encourage improvement in the quality of learning, following what is mandated in Law Number 20 the Year 2003 concerning the National Education System

(Nurrohman, 2018). Therefore, evaluating the implementation of educational assessments is an inseparable part of the Educational Assessment Standards so that these minimum standards can always be improved from time to time to keep abreast of developments in science and technology.

Tutors involved in teaching suggest that current teaching becomes more complicated as a result of an increasingly demanding curriculum and increasing diversity among students and this situation suggests the need for changes in tutor preparation programs that enable them to become more practical in terms of assessment (Goedhart et al., 2019). To prepare quality tutors, assessment practices need to be monitored appropriately and carefully evaluated to develop the tutor's abilities. The assessment methods that have been used must support the development of tutors when they carry out teaching practices (Yan & Cheng, 2015). Teaching practice evaluation methods can be divided into two broad categories, namely traditional assessment methods and authentic assessment methods in the teaching context available to assess teaching practices among current tutors.

However, the tutors' abilities are still far from standard abilities. Low learning outcomes have been a scourge on Indonesia's education system and may undermine the country's future competitiveness (Harjanto et al., 2018). Problems in improving the quality of non-formal education are influenced by several factors; one of the main factors is the quality of Educators and Non-Formal Education Personnel in this case relating to educational qualifications and competence of educators/tutors. The general problem faced is the academic qualifications of about 40% of 121,301 educators and education staff have not met the minimum qualifications following National Education Standards. Besides, the competence of Educators and Education Personnel reaches 60% of 121,301 people working not according to their expertise (miss-match), meaning that it has not been fulfilled according to the ideal expectations demanded by program organizers, even the implementation of professional certification for Non-Formal Education Educators and Education Personnel (Sutisna, 2015).

Capability Assessment of tutors which consists of planning, implementation, and reporting, is still poor because the fact is based on the observation. Leonard (2016) states that the quality of teachers in Indonesia is still below standard. In the 2015, the results of teacher competency tests were still below the minimum competency standard (SKM). At that time, the average score

was 53 or two points below SKM. In terms of learning and assessment methods, a large majority of the teachers in Indonesia are not yet capable of adjusting to the times. A large majority of the teachers are limited to teaching. This method is an old method. Because in the past the teacher was the only source of knowledge. From the UKG test, only 6 percent of teachers can apply assessments well.

So, it is necessary to increase the ability of tutors as teachers in non-formal education. Improving tutor quality is a crucial step for improving PKBM quality, few studies have been undertaken in Indonesia to investigate how tutor engage their students in learning processes in the classrooms, but many tutors do rote-learning seem to be the most dominant mode of teaching (Fahmi et al., 2011). So, the tutor's assessment of students is not optimal until the analysis and evaluation stage.

Therefore, this study is intended to assess tutor reaction patterns on classroom assessment practice skills. The assessment practice carried out by tutors is an important element in addressing student learning needs and ultimately can improve the education system and accountability. Understanding tutor practice assessment serves as a way to find out whether tutors determine or use quality assessment methods that can address student learning needs.

The explanation above indicates that there is a need to identify competencies, backgrounds, and experiences of academic tutors based on the quality of learning that occurs. Being given the importance of assessment for the success of information related to learning including in science learning as well as paying attention to various assessment techniques, the frequency of use of assessment techniques by science tutors is important for research. Information about the frequency of use of assessment techniques can be made to attend ideas or trends in assessment techniques used mainly in informal learning. The tendency of using spectrum in terms of assessment techniques that have long been taught can also be a picture of professionalism of tutors.

METHODS

This research is a descriptive exploratory study that describes and reveals the implementation of assessment to find out the characteristics of tutors' characteristics in the use of assessment techniques in learning science in non-formal schools and illustrates their tendencies. The use

of assessment techniques that focus on teaching terms that reflect tutors' teaching experience, participation in training assessments, and the level of tutor education. This research uses a descriptive quantitative approach with survey methods (Salaria et al., 2012). The data collection technique uses purposive sampling with the aim of tutors who have taught at PKBM for 5 years. The population in this study were PKBM tutors in Jakarta at Package A, Package B, and Package C, namely 158 tutors with details as follows Package A 45 tutors, Package B 53 Tutors, and Package C 60 tutors. Data were collected using a questionnaire, and Focus Group Discussion (Salamah, 2018). Questionnaire data and FGD (Forum Group Discussion) complement each other in which several things are not visible in the FGD can be studied in depth with a questionnaire database, as well as various questionnaire data can be interpreted more deeply with the FGD database.

The scope to be explored in this research activity is (1) assessment techniques and instruments (including attitude competencies, knowledge competencies, and skills competencies); (2) assessment mechanisms and procedures carried out by educators and education units; (3) implementation and reporting of assessments conducted also by educators and education units. According to Miles & Huberman (1992), the collected data is then analyzed using quantitative or qualitative descriptive approaches developed. The stages of qualitative data analysis are data collection, reduction, display, and conclusions. The analysis process starts from the data recap, data reduction includes data simplification by sorting out the data needed. The results of the reduction data are classified according to the analysis design that has been designed which is then displayed. Each reduction has been classified as verified by various field facts, including validation results and student achievement test results. After the data display has been verified, a conclusion is drawn.

The data analysis was conducted using Confirmatory Factor Analysis or CFA to see the attachment between the dimensions measured. Attachment of Capability Assessment dimension to Assessment implementation. CFA intentionally based on theories and concepts, to obtain new variables or factors that represent several items or sub-variables, which are observable variables or observer variables. The main purpose of factor analysis is to explain the structure of relationships between many variables in the form of factors or latent variables or formed variables.

RESULTS AND DISCUSSION

Assessment can describe techniques and problems that educators must consider when they design and use assessments. The characteristic of assessment influences what is learned and the level of meaningful involvement by students in the learning process. The ability to carry out assessments based on length of teaching. The experience in teaching tutor can affect the competence of ratings as shown in table 1 as follows.

Table 1. Capability Assessment Grade

Scale	≤ 5	6-10	>10
Planning	21	28	27
Implementation	19	23	26
Reporting	23	27	28
Capability Assessment	63	78	81

Descriptive statistical results show that tutors' competence in conducting an average rating of 63 points for tutors <5 years 78 in the range of 6-10 and 81 in > 10 years means that the tendency for assessment competencies is high between 6-10 and very high in > 10 years. Old teaching experience shows the score of the tendency for teaching tutors to get higher. Darling-Hammond et al. (2013) emphasizes that tutors' experience in teaching can influence tutors' competence in controlling learning activities in each class. Tutor quality indicators make statistically significant contributions to student outcomes.

Hughes et al. (2018) describe that tutors' competencies in assessment center on three aspects: tutors as learners, tutors learn to learn, and tutors develop the ability to take the knowledge and educate their students. Considering these three factors, it is important to consider constant professional development (whether personal development or professional development that needs to occur to keep pace with today's changing technology). Tutors participating in professional development have varying degrees of comfort, knowledge, and skills in terms of manufacturing space and pedagogical approaches and tools that can be applied in it.

Also, class assessments are carried out through various techniques. The technique is divided into two groups, namely cognitive tests and performance tests. The test technique consists of two types namely essay test and a multiple-choice test. In practice, the forms often used are multiple-choice and free descriptions. Whereas the per-

formance appraisal can be done such as performance appraisal, attitude appraisal, and project appraisal. In the assessment of the knowledge, the Tutor uses more test techniques both written in the form of multiple-choice and description, for the assessment of attitudes usually observation or self-assessment, for assessment of tutor skills using performance appraisal and project appraisal. These findings are illustrated in table 1.

Table 2. Assessment Implementation

Scale	F	%	SD
Practices assessment	38	24.05	0.85
Objective test	54	34.18	0.74
Use of assessment essays	66	41.77	0.65

Data in table 2 shows that most tutors use assessment essays. Essay assessment is used by tutors because essay assessment as one form of competency-based assessment is demanding students to organize, formulate, and express their answers. Learners do not choose answers, but instead give answers in their own words freely based on their understanding. Essay assessment can be carried out in an integrated manner with the learning process and ongoing. Also, compared to multiple-choice, tutors make an essay assessment because of the time and low cost of using it.

Essay assessment according to Singh et al. (2013) can be used as student feedback so that they can self-assess their work and improve it. The results of student essay work are very important as an input to the tutor about student abilities. In this case, the tutor can also find out the problems of assessment in the classroom, among others, the tutor can also find forms of questions that are effective as well as those that are less effective to be applied in his class. Also, tutors can know for certain what learning activities are less successful.

Also, the second assessment used by tutors in PKBM is an objective test. An objective test is a type of learning achievement test consisting of items that can be answered by the learners by choosing one of the answers among several possible answers. Objective tests include objective test kits. Other objective test kits that can be used in evaluating learning outcomes include: 1) True-False objective tests; 2) Matching Tests; 3) Completion Test of objective forms; 4) Fill in Test objective forms. In ideal situations, objective tests can reflect student learning outcomes from cogni-

tive aspects only. Note, only the cognitive aspect. Only one aspect of learning outcomes is evaluated on students. That is, in plain view, maybe we can already see the level of student mastery of teaching material.

The reason why tutors use multiple-choice is that it is easy to correct. But tutors must be wiser not only to use this type of test all the time. Demanding PG test, tutors cannot facilitate the process of developing thinking skills and solving problems optimally. Students determine the choice of answers not because of understanding, but rather than not filled. They become lazy to think, fickle, doubtful, and never mature in making decisions.

Table 3. Written Test

Written Test	Tutor Experienced		
	≤ 5	6-10>	10
Essays	0	13	5
Multiple Choice	9	7	9
True-False	0	1	1
Match	1	2	0
Description	16	17	10

From table 3, the frequency of the range of usage tests was written to the teaching experience of tutors from three schools ≤ 5 years using multiple-choice tests with an average of 9 times, improving with an average of 1 time, and description with an average of 16 times. For a long time teaching 6-10 years on a short field test obtained an average of 13 times, multiple-choice obtained an average of 7 times, true-false obtained an average of 1 time, asking for an average obtained twice, average description obtained 17 times. And for a long time teaching for more than 10 years using a short field obtained an average of 5 times, multiple-choice obtained an average of 9 times, true-false obtained an average of 2 times and the description obtained an average of as much as 10 times. These results indicate that the tutor group with 6-10 years of teaching experience uses the written test format most often and is the most diverse compared to other groups. The average is obtained from several written test formats using science tutors for one semester.

These data indicate that classroom assessment becomes an important component in effective learning that is carried out based on the assessment potential. This study revealed that tutors used oral questions to assess students in science classes. This is because written tests have

the advantage of being able to measure the ability or behavior of students in a goal, can measure the ability to organize ideas, and students can explore answers using their language. The frequency of using cognitive assessment techniques in learning science-based on tutors' long teaching experiences is seen between 6-10 years more dominant using various cognitive assessment techniques. It was also revealed that the teacher did not use different techniques such as group work, written assignments, practical work, or investigative work to assess students.

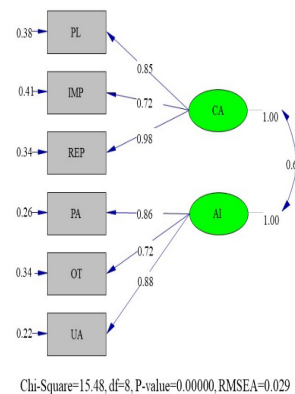


Figure 1. Capability and Implementation Assessment Analysis

Figure 1 shows the CFA measurement model which has 2 constructs with 6 manifest dimensions for the capability assessment. There are 3 dimensions and the manifest implementation assessment has 3 dimensions. The loading factor value is in the middle of the direct line towards the manifest variable. Meanwhile, the score that is behind the manifest variable is the standard error of each manifest variable.

The structural model characterizes the relationships between the constructs or the latent variables and defines those latent variables that indirectly or directly cause alterations in the values of other latent variables in the model (Byrne, 2013). A set of CFA was developed as part of the analysis. The previously presented Initial Capability assessment (see Figure 1) is the initial path of the theoretical structure model. The outcome of CFA revealed that the χ^2 of 15.48 with 8 df was statistically significant with $p < 0.05$, which suggested an inappropriate fit. Byrne (2013) noted that the χ^2 is highly sensitive to sample size and frequently recommends a poor fit with a large sample size. The remaining fit statistics revealed a moderate fit between the theoretical model and the data with slightly lower than the common-

ly accepted scores of 0.50 ($\chi^2/df = 1.94$; RMSEA=0.02). Figure 1 shows that there is a link between appraisal capabilities and implementation. seen from the p-value of 0.64 or greater than 0.5.

Merchie et al. (2018) state that the work of tutors in carrying out their duties as educators in certain units is following the duty letter from the authorized institution (local government or education provider groups). Tutor experience is also related to tutors' competencies supported by Minister of National Education regulation 16 of 2007 concerning academic qualification standards and tutor competencies on four competency standards, namely pedagogical competence, tutors' ability to manage learners' learning, including mastery of student characteristics, learning theory, learning organization, the benefits of information technology, facilitating the development of potential learners, being able to communicate effectively, being able to assess and evaluate learning outcomes obtained, and reflecting on improving the quality of learning (Markiewicz & Patrick, 2016).

From the data taken tutors to do activities in general, namely: evaluating learning material, presenting daily, providing practical guidance in mastering the material, providing feedback, and correction. So, teaching tutors can increase the tutors' potential intelligence and professionalism, and can also influence the assessment techniques used by tutors. There are several tutor grading techniques used, one of which is a written test where the test is a written test to be answered by students by giving written answers. The choice of test form can be determined by the purpose of the test, the number of test-takers, the time available to check the students' answer sheets, the scope of the material, and the subject to be tested. Thus, the test given by the tutor depends on the suitability of the technique with the conditions of students in learning.

Previous research has suggested that tutors with teaching experience are more accurate in assessing students' understanding of being people who share their experiences with other tutors (D'Mello & Graesser, 2012). Therefore, comparing assessments conducted by tutors needs to be done so that they become material for identifying tutor research competencies. To achieve these objectives an approach that is needed; the first approach is studying the effect of students' general characteristics and the second approach is studying the structure and learning processes of students to study learning outcomes. Graesser (2011) in his study found that tutors will have

more information about students who are judged than using only one approach, so it is recommended that tutors continue to explore their ability to judge.

Tutors as professional educators are not only required to be able to convey information to students but are also asked to plan, manage, diagnose, assess the process, and the results of the teaching and learning process. The characteristics of tutors can be seen from the effectiveness of tutors in teaching. Time spent by tutors on academic activities is one of the characteristics of an effective tutor and added to the effectiveness of tutors marked by six tutors to carry out activities, namely: carrying out daily evaluations, presenting learning materials, providing practical guidance for mastering the material, providing feedback and correction (Mertens, 2014). Avalos (2011) also suggests that the characteristics of effective learning conducted by tutors are giving attention to students on teaching materials, giving sufficient time for students to understand the material, mastering the material and abilities needed in teaching, and always monitoring the progress of learning. Also, the tutor's background and experience may influence effectiveness and learning performance.

The teaching practice of teacher trust must be linked in a meaningful way, as do other teaching assignments, such as student assessment, where the conception of assessment and assessment practice are significantly related. The reason tutors adopted class management practices such as the use of prizes and penalties has been analyzed. Tutors' main argument for using such practices is their efficiency: They offer easy and easy-to-use solutions for reacting to wrong behavior (Muñoz et al., 2012). Besides, this gives the tutor the feeling of keeping students under control. At the other end of the spectrum, practices such as personal praise and verbal encouragement are believed to be of little use because tutors say that such practices cannot work with students' tutors.

Based on the results of the study, it can be concluded that using an assessment strategy by the tutor is not appropriate to develop students' problem-solving skills and does not consider the assessment of students' cognitive style of science in applying the assessment strategy. Nevertheless, each student has the characteristics of an individual cognitive style. In the tutor experience that is widely used in non-formal education in Indonesia and elsewhere to evaluate the quality of education, the tutor's experience that most clearly distinguishes the best and worst courses and cor-

relates most closely with student performance is "Teaching staff here usually provide useful feedback about how they work in assessing students (Van den Bos & Brouwer, 2014).

This does not mean that tertiary tutors provide useful feedback - it means whether they provide useful feedback or not make a greater difference than anything they do. There is a growing consensus that more attention must be paid to the personal relevance of what is taught in programmed tutor education and that the experience of novice tutors in teaching practice must be taken into account.

Offerdahl & Tomanek (2011) explain how changes in individual thinking about assessment may not lead to changes in teaching practice. In this case study, three educators consider student center strategies, such as using a formative assessment that reveals information about student progress to inform teaching. These strategies are implemented but ultimately adapted to previous didactic approaches that focus on providing students with the correct answers. Offerdahl & Tomanek (2011) speculate that the level of dissatisfaction is stronger with the need for assessments to stimulate change.

Their study highlights three problems. First, it is more difficult to change assessment practices than to change theoretical understanding. Second, there is little data that reveals the reasons for educator assessment choices. Finally, if institutions and departments want to support individual educators to improve assessment, they must consider the influence of many of the contextual factors that shape the practice of educators. Describing positive or negative ways in which formative assessment can influence student retention and emphasize its role in academic integration. Conversely, the feedback itself is more likely to be regarded as a comment about what has been learned. In the absence of signs, it has been reported that students read feedback more carefully and use it to guide their learning. Being given the evidence of this research (school-based) some schools have adopted a policy that all assignments only have feedback and no sign must be given. The 'College as a learning' grading system in colleges is perhaps the best-known example of higher education from 'less-grade' assessments.

Assessment sometimes seems, at the same time, very expensive, not liked by students and tutors, and is largely ineffective in supporting learning. Concerning these problems, the remainder

of this article describes and seeks to justify a series of 'conditions in which assessment can support learning'. Evidence is rarely conclusive enough to argue that if your assessment meets these requirements, learning will be more effective. They are offered as a set of reasonable guidelines. In this study, the results show that the preparation of assessment in non-formal education, there is an influence between the preparations of the assessment of the implementation of the assessment. So, it is necessary to improve tutor assessment preparation.

Although this research reveals some important findings that can inform policy and practice, the findings are only limited to one type of research method. Focusing only on perception surveys without observing actual assessment practices, analyzing relevant documents, and dialoguing with tutors are the main limitations of this study. A qualitative approach might have provided a clearer picture of assessment practices in the class that is considered skilled tutors and highlighted which they use more often and why. Therefore, further research where a qualitative approach is used is recommended. Students' needs for tutor support change as they become more cognitively engaged and develop expertise. Tutor needs to measure how much scaffolding to provide as individual learners become more knowledgeable and proficient. However, at any stage of development, learners benefit from strategically placed direct instruction, feedback, and critical questions that guide their learning.

CONCLUSION

The results of the study showed that teaching tutors' competencies of less than 5 years had minimum abilities compared to teaching tutors for more than 10 years, which meant that the competency of teaching experience assessment determined tutors' teaching competencies. In addition, most tutors use essay grading. Essay grading is used by tutors because essay grading is one form of competency-based assessment that requires students to organize, formulate, and express their answers. The reason why tutors use multiple choice is because it is easy to correct. But tutors need to be wiser not just to use this type of test all the time. Tutors complain that multiple choice tests cannot facilitate the process of developing thinking skills. In this finding, tutors' experience becomes a matter that needs to be widely used in non-formal education to improve the quality

of education, the tutors' experience most clearly distinguishes the good and bad performance of students. But what needs to be emphasized is the collaboration between experienced tutors and young tutors so that optimal learning can be established.

REFERENCES

- Avalos, B. (2011). Teacher professional development in teaching and teacher education over ten years. *Teaching and teacher education*, 27(1), 10-20.
- Blazar, D., & Kraft, M. A. (2017). Teacher and teaching effects on students' attitudes and behaviors. *Educational evaluation and policy analysis*, 39(1), 146-170.
- Borg, S. (2018). Evaluating the impact of professional development. *RELC Journal*, 49(2), 195-216.
- Brandt, J. O., Bürgener, L., Barth, M., & Redman, A. (2019). Becoming a competent teacher in education for sustainable development. *International Journal of Sustainability in Higher Education*, 20(4), 630-653.
- Byrne, B. M. (2013). *Structural equation modeling with Mplus: Basic concepts, applications, and programming*. routledge.
- Copriady, J. (2014). Teachers Competency in the Teaching and learning of Chemistry Practical. *Mediterranean Journal of Social Sciences*, 5(8), 312-312.
- Darling-Hammond, L., Newton, S. P., & Wei, R. C. (2013). Developing and assessing beginning teacher effectiveness: The potential of performance assessments. *Educational Assessment, Evaluation and Accountability*, 25(3), 179-204.
- Darling-Hammond, L., Flook, L., Cook-Harvey, C., Barron, B., & Osher, D. (2020). Implications for educational practice of the science of learning and development. *Applied Developmental Science*, 24(2), 97-140.
- De Vries, S., van de Grift, W. J., & Jansen, E. P. (2014). How teachers' beliefs about learning and teaching relate to their continuing professional development. *Teachers and Teaching*, 20(3), 338-357.
- D'Mello, S., & Graesser, A. (2012). Dynamics of affective states during complex learning. *Learning and Instruction*, 22(2), 145-157.
- Fahmi, M., Maulana, A., & Yusuf, A. A. (2011). Teacher certification in Indonesia: A confusion of means and ends. *Center for Economics and Development Studies (CEDS) Padjadjaran University*, 3(1), 1-18.
- Goedhart, N. S., Blignaut-van Westrhenen, N., Moser, C., & Zweckhorst, M. B. M. (2019). The flipped classroom: supporting a diverse group of students in their learning. *Learning Environments Research*, 22(2), 297-310.
- Graesser, A. C. (2011). Learning, thinking, and emoting with discourse technologies. *American Psychologist*, 66(8), 746.
- Harjanto, I., Lie, A., Wihardini, D., Pryor, L., & Wilson, M. (2018). Community-based teacher professional development in remote areas in Indonesia. *Journal of Education for teaching*, 44(2), 212-231.
- Hooper, L., Begg, M. D., & Sullivan, L. M. (2014). On Academics Integrating Competencies and Learning Outcomes in Core Courses for the MPH. *Public Health Reports*, 129(4), 376-381.
- Hughes, J., Morrison, L., & Dobos, L. (2018). Re-Making Teacher Professional Development. *Studies in health technology and informatics*, 256, 602-608.
- Koloi-Keaikitse, S. (2016). Assessment training: A precondition for teachers' competencies and use of classroom assessment practices. *International Journal of Training and Development*, 20(2), 107-123.
- Leonard, L. (2016). Kompetensi tenaga pendidik di Indonesia: Analisis dampak rendahnya kualitas SDM guru dan solusi perbaikannya. *Formatif: Jurnal Ilmiah Pendidikan MIPA*, 5(3).
- Markiewicz, A., & Patrick, I. (2016). *Developing monitoring and evaluation frameworks* (No. BOOK). Sage Publications.
- Merchie, E., Tuytens, M., Devos, G., & Vanderlinde, R. (2018). Evaluating teachers' professional development initiatives: towards an extended evaluative framework. *Research papers in education*, 33(2), 143-168.
- Mertens, D. M. (2014). *Research and evaluation in education and psychology: Integrating diversity with quantitative, qualitative, and mixed methods*. Sage publications.
- Miles, M. B., & Huberman, A. M. (1992). Analisis Data Kualitatif, Terjemahan Tjejep Rohendi. Jakarta: Universitas Indonesia.
- Muñoz, A. P., Palacio, M., & Escobar, L. (2012). Teachers' beliefs about assessment in an EFL context in Colombia. *Profile Issues in Teachers Professional Development*, 14(1), 143-158.
- Nurrohman, M. H. (2018). *Telaah isi UU RI No. 20 tahun 2003 Pasal 03 dan Relevansinya dengan Pendidikan Karakter pada Sekolah Dasar* (Doctoral dissertation, IAIN Ponorogo).
- Offerdahl, E. G., & Tomanek, D. (2011). Changes in instructors' assessment thinking related to experimentation with new strategies. *Assessment & Evaluation in Higher Education*, 36(7), 781-795.
- Raharjo, S. B. (2012). Evaluasi trend kualitas pendidikan di indonesia. *Jurnal Penelitian dan Evaluasi Pendidikan*, 16(2), 511-532.
- Salamah, U. (2018). Penjaminan Mutu Penilaian Pendidikan. *EVALUASI: Jurnal Manajemen Pendidikan Islam*, 2(1), 274-293.
- Salaria, N. (2012). Meaning of the term descriptive survey research method. *International journal of transformations in business management*, 1(6), 1-7.
- Sandi, P. A., Sutarto, J., & Yusuf, A. (2018). Home-schooling Model Management (A Study at Anugrah Bangsa Package A Homeschooling Semarang). *Journal of Primary Education*, 7(2), 204-210.

- Singh, T., Gupta, P., & Singh, D. (2013). *Principles of medical education*. Jaypee Brothers Medical Publishers.
- Sutisna, A. (2015). Pengembangan Model Bimbingan Teknis Berkelanjutan dalam Meningkatkan Kompetensi Tutor Paket C. *Visi*, 10(2).
- Van den Bos, P., & Brouwer, J. (2014). Learning to teach in higher education: how to link theory and practice. *Teaching in Higher Education*, 19(7), 772-786.
- Yan, Z., & Cheng, E. C. K. (2015). Primary teachers' attitudes, intentions and practices regarding formative assessment. *Teaching and Teacher Education*, 45, 128-136.