JPPIPA 8(3) (2022)

Jurnal Penelitian Pendidikan IPA

Journal of Research in Science Education

http://jppipa.unram.ac.id/index.php/jppipa/index

Construction and Validation of Evaluation Instruments for Science Learning Programs Based on Context, Input, Process, And Product (CIPP) Models

Cindy Alvianita[1], Tanti[2*], Bambang Hariyadi[1]

[1] Pascasarjana Pendidikan IPA, Universitas Jambi, Jambi, Indonesia.
[2] Universitas Islam Negeri Sulthan Thaha Saifuddin Jambi, Jambi, Indonesia.

DOI: 10.29303/jppipa.v8i3.1369

Article Info
Received: February 1, 2022
Revised: June 18, 2022
Accepted: July 20, 2022
Published: July 31, 2022

Abstract: This study aims to develop an evaluation instrument for science learning program in junior high schools based on the CIPP (Context, Input, Process, Product) evaluation model. This research uses the exploratory sequential mix method for instrument development. The qualitative stage was carried out through in-depth interviews with five principals and ten teachers. The qualitative stage aims to analyze needs in the field and explore the opinions of school principals and teachers regarding the evaluation of science learning programs. Researchers distributed instruments to 112 participants for the quantitative stage of the validation test. The instrument was validated in three stages. The first stage is the face validity test, the second stage is the Exploratory Factor Analysis (EFA) test, and the last stage is the Confirmatory Factor Analysis (CFA) test. The results of face validity conducted with two experts stated that each indicator developed had followed the flow of the Context, Input, Process, and Product (CIPP) model. Exploratory Factor Analysis Test Based on scree plot data, it was found that there was a fault after the four components, and the researcher decided to take four elements with a total variance of 50.95%. The confirmatory Factor Analysis (CFA) test showed that from 47 items, only 35 items were declared valid. This research has positive implications for the practice of integrated science learning at the junior high school level in Indonesia. Principals and teachers can use this CIPP-based evaluation instrument to evaluate the effectiveness of the integrated science learning process that has been carried out so far.

Keywords: CIPP Model; CFA; Evaluation; EFA; Factor Analysis

Citation: Alvianita, C., Tanti, T., & Hariyadi, B. (2022). Construction and Validation of Evaluation Instruments for Science Learning Programs Based on Context, Input, Process, And Product (CIPP) Models. *Jurnal Penelitian Pendidikan IPA*, *8*(3), 1089–1095. https://doi.org/10.29303/jppipa.v8i3.1369

Introduction

Learning natural Sciences at the Junior High School level in Indonesia is carried out in an integrated manner (integrative science). According to Wilujeng (2018), integrative science is a science learning approach that connects or integrates various fields of science studies into one unified discussion Substantively, integrative science is a tool to develop students' cognitive, affective, psychomotor, and creative abilities. However, the low achievement of Indonesian students in international events that measure scientific literacy, such as PISA, shows the low quality of science learning in Indonesia. Based on the OECD report, Indonesian students' average scientific literacy score on the 2018 PISA test was 396. This score is still far below the OECD average of 489 (Schleicher, 2019).

A review of various research literature shows that students still experience various obstacles in learning science (Osborne et al., 2003; Potvin & Hasni, 2014).

* Corresponding Author: tanti@uinjambi.ac.id

Copyright © 2022, Author et al.
This open access article is distributed under a (CC-BY License

Research conducted by Potvin and Hasni (2014) shows that one of the obstacles faced in learning science is students' low motivation and attitude towards science. Even Osborne et al. (2003) emphasize a phenomenon called "swing away from science," which describes students' low interest and motivation in studying science, especially physics. Osborne et al. (2003) observed several factors that led to the emergence of this phenomenon, one of which was the teacher factor. According to Osborne et al. (2003), the ability of teachers to be able to contextualize science material following the daily experiences of students, clarity in communicating the learning objectives to be achieved, and the ability of teachers to be able to facilitate various types of cognitive styles of students are factors that influence student success in learning science. Dari et al. (2022) research shows that students' low motivation in studying physics has an impact on the low cognitive abilities of students. Furthermore, Dari et al. (2022) state that teachers' use of conventional learning strategies leads students to memorize various facts rather than understand concepts. Another study was conducted by Rezeki et al. (2021), which showed that low concept understanding was one of the obstacles for students in learning science. The findings of this study are in line with the initial observations made by the researcher in one of the junior high schools in Jambi City. Researchers found various obstacles students face in science learning, one of which was inappropriate learning media. When the teacher provides learning media in the form of a video, the teacher does not re-explain the purpose of the video so that students find it challenging to understand the material. In addition, most of the students are passive during the learning process. The main activity of the students is just listening to the explanation from the teacher.

The problems that arise in integrated science learning raise questions regarding the effectiveness of the integrated science learning programs that have been implemented so far. Therefore, to assess the learning program's effectiveness, it is necessary to evaluate it. According to Stronge and Tucker (2020), evaluation is an assessment process to analyze the differences in achievement with specific standards and how the benefits that have been carried out are compared to the expectations obtained. Evaluation of science learning programs can be said to be a process of collecting scientific data or information related to the science learning process itself and diagnosing strengths and weaknesses to analyze the efficiency of the implementation of science learning programs. The principal evaluates the implementation of the Integrated Science learning program as top management. As Darma (2019) revealed, evaluation of learning programs must be carried out on an ongoing basis. Furthermore, Darma (2019) stated that the assessment results would be feedback for the top education management to improve teacher performance and professionalism. By evaluating the integrated science learning program, it is hoped that the quality of learning will be better. In the end, it will grow student demand and motivation to learn. The interview results that the researcher conducted with one of the teachers stated that so far, the principal only evaluated the lesson plan made by the teacher. A thorough evaluation of the effectiveness of the implementation of integrated science learning is rarely carried out. One of the reasons is the absence of proper evaluation instruments. A good and thorough evaluation will determine what is desired from teaching and learning activities and be one of the strategies that need to be done to find out whether the teaching and learning process in schools has achieved the expected results. Therefore, it is necessary to develop an appropriate learning program evaluation instrument that needs to be designed so that school principals can evaluate teachers implementing science learning.

In this study, researchers used the Context, Input, Process, and Product (CIIPP) evaluation model developed by Stufflebeam in 1967 (Hakan & Seval, 2011). This model aims to complement the basis of decisions in system evaluation with analysis oriented to planned changes. Aziz, Mahmood, and Rehman (2018) states that Stufflebeam developed the CIPP Evaluation Model in 1967 to evaluate the effectiveness of programs or projects from various disciplines such as education, transportation, military, community development, and transportation. Compared with other evaluation models, this model is complete because it includes formative and summative aspects. The main characteristics of the CIPP model are the four components of evaluation – context, input, process, and product. These four evaluation dimensions form the basis for planning, structuring, implementation, and recycling decisions. Context evaluation attempts to describe and detail the environment, unmet needs, population and sample served, and project objectives. The context evaluation is carried out to answer the questions: (1) What program activities have not met needs; (2) What development goals are related to meeting needs; (3) What are the most easily attainable goals? In short, it can be said that context evaluation is an evaluation of the circumstances surrounding learning.

Input Evaluation helps regulate decisions, determine available sources, what alternatives are taken, what plans and strategies are to achieve goals. This evaluation helps regulate decisions, determine the available resources, what alternatives are taken, what plans and strategies are to achieve the needs, and the work procedures to achieve them. Meanwhile, evaluation of the process component includes collecting data obtained when designing and running the program. In evaluating the learning program, the process evaluation assesses whether the implementation of teaching and learning activities follows the learning objectives to be achieved. Questions related to evaluating process components include: is the learning program implemented properly? What are the factors that hinder the success of the program? What important changes need to be made? The answers obtained at the process evaluation stage will assist the principal in controlling and directing performance in implementing science learning programs in schools. In addition, process evaluation is carried out to detect or predict problems that may be encountered in learning activities. Product Evaluation is an evaluation carried out to measure the program's results being implemented. This type of evaluation is carried out to assess whether educational activities are suitable or not. In other words, product evaluation is carried out to measure the relevance of the output to the context, input, and process of a learning program.

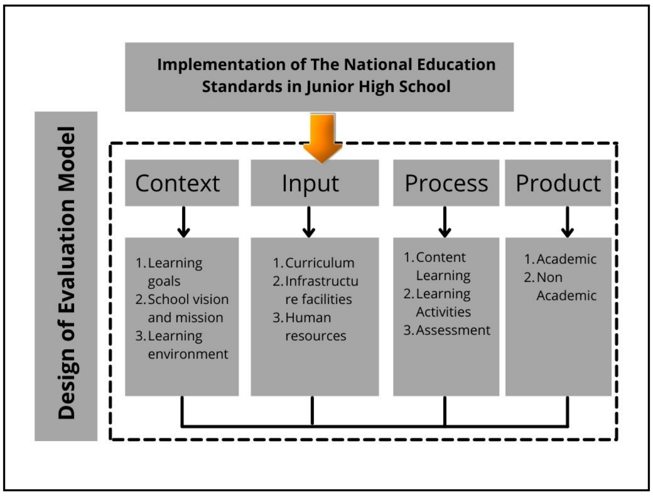

Figure 1. The Design of the CIPP model adaptation from Darma (2019)

Method

This study uses a mixed-method research design. According to Creswell and Clark (2017), a mixed-method is a procedure for collecting, analyzing, and mixing quantitative and qualitative data through several stages of research in a single study to understand the problem in more depth. The design used is an exploratory mixed-method design for instrument development. The flow of research design using mixed exploration methods for instrument development can be seen in the Figure 2.

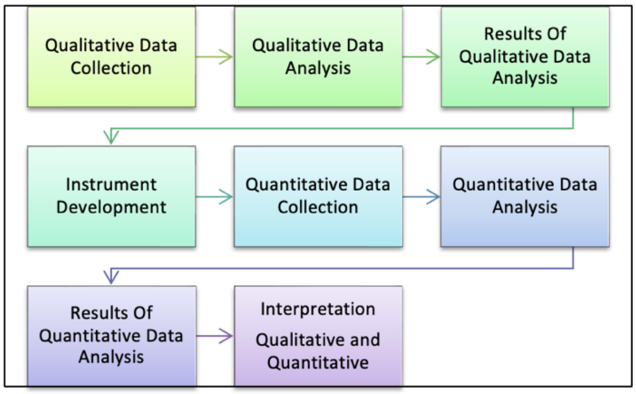

Figure 2. Exploratory Mixed Method Design

The initial stage of this research is the collection of qualitative data. At this stage, the researcher analyzes various references to explore the theoretical basis of the CIPP model concept and its implementation to evaluate the learning program. Analysis of qualitative data include selecting noteworthy quotations, classifying them with pertinent subjects, and developing larger themes (Mihas & Institute, 2019). In this research, the researchers conducted in-depth interviews with ten Integrated Science teachers and five principals from five junior high schools in Jambi city to find out their opinions regarding the implementation of the learning evaluation that had been carried out so far. In-depth interviews were also conducted to explore the views of teachers and school principals regarding what factors should be evaluated in Integrated Science learning. The researcher's data from the literature study and interviews were then reduced. Researchers choose which data is most relevant to be used to support research. From these data, then it is compiled into an instrument for further validation with experts. Validation is carried out by two experts, namely material and design experts. Material expert validation was carried out to determine whether each indicator that the researcher compiled was under every aspect of learning required. Design validation is carried out to determine whether the product design, in this case, the product display is as expected. Design validation also concerns the use of grammar, punctuation, and the suitability of letters under the target user of the instrument later.

The second stage of research is the quantitative stage. The draft evaluation instrument that the expert panel has validated is then distributed to 112 respondents consisting of science teachers and school principals in Jambi Province for validation. Quantitative validation uses two stages, namely exploratory factor analysis using SPSS and confirmatory factor analysis (sem pls). Several assessment indicators on factor analysis include measurement of Kaiser Meyer Olkin (KMO) value, Bartlett's value, loading factor, eigenvalue, scree plot, and rotation of Oblimin with

Kaiser Normalization. Furthermore, the CFA test was carried out using variance-based structural equation modeling (SEM with PLS). SEM is a multivariant analysis that combines factor analysis with path analysis simultaneously. SEM model is an analysis that integrates empirical data analysis with theoretical construction. The researcher uses the SEM model because the relationship between variables is very complex, but the data sample size is small. This aligns with Haryono (2017) opinion, which states that the SEM with the PLS test can be carried out on complex modeling analyses with relatively small samples.

Result and Discussion

From the literature study, several evaluation items of the learning program were arranged. This evaluation item then became a reference for conducting interviews with three science teachers and two school principals. Based on the interview results, several additional evaluation items were obtained. Other evaluation items include implementing blended learning-based learning, namely face-to-face and online learning, to increase some learning needs. According to participants, advances in information technology coupled with the Covid-19 pandemic have changed education. These include learning resource needs that must be diverse, meeting face-to-face learning and online learning, using virtual laboratories, and using appropriate learning management systems. Participants' statements were then used as additional items in the developed evaluation instrument. The initial draft of this evaluation instrument was then tested for validity by a panel of experts, namely material experts and design experts. Based on the validation test conducted by the expert panel, it was found that the initial draft of the evaluation instrument developed was under the flow of the CIPP model. In addition, the expert panel suggested that the ambiguous sentence in item 35 be revised.

The exploratory factor analysis test aims to reduce and categorize some statement items into new dimensions (Pallant, 2010). In this study, a factor analysis test was conducted to determine whether the measurement construct of the CIPP model evaluation instrument consists of four factors (context, input, process, and product). The initial assumption test shows the Kaiser-Meyer-Olkin (KMO) value of 0.795 and the value of Bartlett's Test of Sphericity of 3.184E3 with a significance value of 0.000. The KMO and Bartlett's Test of Sphericity values indicates that they have met the requirements for a factor analysis test (Pallant, 2013), as shown in Table 1 below:

Tabel 1. KMO and Bartlett's test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.	0.795
Approx. Chi-Square	3.184E3
df	1081
Sig.	0.000

The results of the principal component analysis with orthogonal rotation (varimax) on 47 items evaluation instruments resulted in eleven components or dimensions with eigenvalues > 1, with a total variance value of 68.933%. However, analysis of the scree plot shows a fracture after four components, as shown in Figure 3 below:

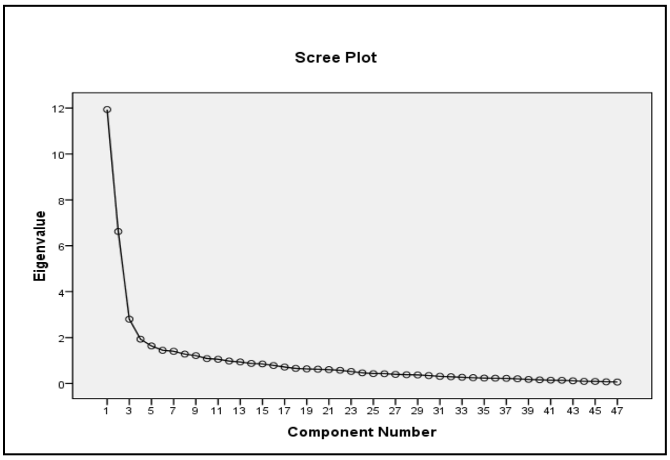

Figure 3. *Screeplot*

Based on the data scree plot above, it can be seen that there is a fault after the four components, and the researcher decided to take four factors. Examination of the communalities shows that several items have values <0.2, namely in item 1 and item 12. According to Cabrera-Nguyen (2010), the value of communalities < 0.3 indicates a low relationship between items and components and must be excluded from the test. The remaining forty-five items will be processed in the re-test of factor analysis. Distribution of the factors according to their names and items are shown in Table 2.

As Table 2 shows, the factor loading of the items ranged from 0.44 to 0.81. The total variance of the four factors was 50.95%. After analyzing the items on the scale, the factors were given meaningful names based on the common characteristics of the items grouped under the factors and skills sought to be assessed. The first factor included twelve items titled "Evaluation of Input," the second factor had fourteen items titled "Evaluation of Context," the third factor had ten items titled "Evaluation of Process," and the fourth factor had nine items titled "Evaluation of Product."

Table 2. Distribution of The Factors

Factors	Factors of CIPP	Items	Eigen Values	% Variance
Factor 1	Evaluation of Input	27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38	11.94	25.40
Factor 2	Evaluation of Context	2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 13, 14, 15, 16	6.62	14.09
Factor 3	Evaluation of process	17, 18, 19, 20, 21, 22, 23, 24, 25, 26	2.80	5.96
Factor 4	Evaluation of product	39, 40, 41, 42, 43, 44, 45, 46, 47	1.93	5.50

The four factors of CIPP instrument identified via exploratory factor analysis was cross-validated by confirmatory factor analysis (CFA) in stage 2, with an independent sample of randomly selected participants. The CFA creates a measurement model for the CIPP instrument evaluation and is a more robust evaluation of underlying factor structure in that items can only load on one factor; loadings on all other factors are set to zero in CFA (Haryono, 2017) The CFA test was carried out using variance-based structural equation modeling (SEM with PLS) in this study. SEM is a multivariate analysis that combines factor analysis with path analysis simultaneously. SEM model is an analysis that integrates empirical data analysis with theoretical construction. Table 3 provides standardized factor loadings and reliability for the final CFA model.

Based on the results of the EFA and CFA tests of forty-seven learning program evaluation items, thirty-five items were valid and reliable. These thirty-five items are grouped into four evaluation dimensions, namely context, input, process, and product, according to the CIPP model. The implementation design of learning, also contained in the context aspect, is a vital component in learning needs. The learning design is incorporated in Regulation of the Minister of Education and Culture number 103 of 2014 that learning implementers must be based on the learning implementation plan. Furthermore, in component 1, the quantitative data results are included in the input aspect. Input evaluation is related to what teachers do to prepare reliable students, one of which is facilitating the various needs of students in the learning process. The indicators in component 1 have followed what is needed in the input aspect. One of the indicators in component 1 is that educators provide opportunities for students to think, analyze, solve problems, and act without fear. This indicator is by the input aspect. According to Anjarsari (2014), thinking, analyzing, solving problems, and acting without fear make students make decisions quickly, precisely, and efficiently. This skill is the provision for students to become reliable students who can compete in the era of globalization.

Table 3. Standardized Factor Loadings and Reliability for the Final CFA Model

Latent Variables and Indicators	Factor Loadings	Cronbach's Alpha	AVE
Factor 1. Evaluation of Input			
IT27	0.73	0.93	0.55
IT28	0.78		
IT29	0.74		
IT30	0.78		
IT31	0.61		
IT32	0.81		
IT33	0.81		
IT34	0.79		
IT35	0.77		
IT36	0.68		
IT37	0.68		
IT38	0.74		
Factor 2. Evaluation of Context			
IT10	0.86	0.79	0.53
IT11	0.79		
IT15	0.56		
IT7	0.64		
IT9	0.76		
Factor 3. Evaluation of Process			
IT17	0.65	0.88	0.51
IT18	0.67		
IT19	0.76		
IT20	0.72		
IT21	0.76		
IT22	0.62		
IT23	0.82		
IT24	0.71		
IT25	0.71		
Factor 3. Evaluation of Process			
IT39	0.76	0.89	0.53
IT40	0.74		
IT41	0.70		
IT42	0.78		
IT43	0.74		
IT44	0.70		
IT45	0.76		
IT46	0.67		
IT47	0.72		

The data results in component 3 are included in the process aspect. According to Pratiwi et al., (2019), the process aspect is related to the teacher's ability to guide the learning process to achieve the learning objectives that have been prepared. One of the things that teachers do in guiding the learning process is indicator 24, facilitating interaction between students and between students and teachers, the environment, and other learning resources, and indicator number 19 asks questions that connect previous knowledge with the material to be studied. The results of the fourth component are included in the product aspect. Product evaluation is an assessment to see the program's success in achieving its goals. Indicators in this aspect only focus on the teacher's evaluation of students. As stated in the Minister of Education and Culture Number 23 of 2016, the assessment aspect of learning must be following the principles of assessment as contained in indicator number 46. The student assessment must also cover three aspects: attitude, knowledge, and skills, as included in indicator number 47. In looking at learning achievement, not only is an assessment carried out in the form of written scores, but activities with students to make lesson summaries can also be a benchmark in measuring learning achievement. This result aligns with Monica and Hadiwinarto (2021) opinion that summarizing lessons can help students prepare themselves for learning, and teachers can know the achievement of the education.

Conclusion

The analysis of the research results shows that the integrated science learning evaluation instrument developed based on the CIPP model (Context, Input, Process, Product) is declared valid and reliable. The CIPP-based integrated science learning evaluation instrument developed by the researcher contained 35 indicators divided into four aspects, namely context (five items), Input (nine items), process (nine items), and product (twelve items). This research has positive implications for the practice of integrated science learning at the junior high school level in Indonesia. The instrument can be used to evaluate the effectiveness of the integrated science learning process that has been carried out so far.

References

Anjarsari, P. (2014). *Pentingnya melatih keterampilan berpikir (thinking skills) dalam pembelajaran IPA SMP.* Paper presented at the Makalah disampaikan dalam PPM "Optimalisasi Implementasi Kurikulum 2013 dengan Workshop Pengembangan LKS IPA Berpendekatan Guided Inquiry untuk Mengembangkan Thinking Skills dan Sikap Ilmiah Siswa.

Aziz, S., Mahmood, M., & Rehman, Z. (2018). Implementation of CIPP Model for Quality Evaluation at School Level: A Case Study. *Journal of Education and Educational Development, 5*(1), 189-206. http://dx.doi.org/10.22555/joeed.v5i1.1553

Creswell, J. W., & Clark, V. L. P. (2017). *Designing and conducting mixed methods research*: Sage publications. Retrieved from https://us.sagepub.com/en-us/nam/designing-and-conducting-mixed-methods-research/book241842

Dari, U., Halim, A., & Ilyas, S. (2022). The Influence of the use of the approach of blended learning model rotation based moodle on motivation and cognitive abilities of students in the subjects of physics. *Jurnal Penelitian Pendidikan IPA, 8*(1), 195-202. https://doi.org/10.29303/jppipa.v8i1.1100

Darma, I. K. (2019). The effectiveness of teaching program of CIPP evaluation model: Department of mechanical engineering, Politeknik Negeri Bali. *International Research Journal of Engineering, IT & Scientific Research, 5*(3). https://doi.org/10.21744/irjeis.v5n3.619

Hakan, K., & Seval, F. (2011). CIPP evaluation model scale: development, reliability and validity. *Procedia-Social and Behavioral Sciences, 15*, 592-599. https://doi.org/10.1016/j.sbspro.2011.03.146

Haryono, S. (2017). *Metode SEM untuk Penelitian Manajemen AMOS Lisrel PLS*. Jakarta: Luxima Metro Media.

Mihas, P., & Institute, O. (2019). *Learn to use an exploratory sequential mixed method design for instrument development*: SAGE Publications, Limited. https://dx.doi.org/10.4135/9781071802779.n18

Monica, S., & Hadiwinarto, H. (2021). Pengaruh Keterampilan Membuka dan Menutup Pembelajaran Terhadap Keaktifan Belajar Siswa di Sekolah Menengah Kejuruan (SMK) Negeri 1 Lubuklinggau. *Ad-Man-Pend: Jurnal Administrasi Manajemen Pendidikan, 3*(2), 12-23. https://doi.org/10.32502/amp.v3i2.3054

Osborne, J., Simon, S., & Collins, S. (2003). Attitudes towards science: A review of the literature and its implications. *International journal of science education, 25*(9), 1049-1079. https://doi.org/10.1080/0950069032000032199

Pallant, J. (2010). *SPSS survaival manual: a step by step guide to data analysis using SPSS*: McGraw-Hill Education.

Potvin, P., & Hasni, A. (2014). Interest, motivation and attitude towards science and technology at K-12 levels: a systematic review of 12 years of educational research. *StudieS in Science education, 50*(1), 85-129. https://doi.org/10.1080/03057267.2014.881626

Pratiwi, M., Ridwan, R., & Waskito, W. (2019). Evaluasi Teaching Factory Model CIPP. *Jurnal Imiah Pendidikan dan Pembelajaran, 3*(3), 414-421. https://doi.org/10.23887/jipp.v3i3.22205

Rezeki, S., Halim, A., Nasrullah, N., Yusrizal, Y., & Ilyas, S. (2021). Implementation of Schoology-Based E-Module Learning Media to Improve Student's Concept Understanding on Temperature and Heat Material. *Jurnal Pendidikan IPA, 7*, 14-19. https://doi.org/10.29303/jppipa.v7iSpecialIssue.887

Schleicher, A. (2019). *PISA 2018: Insights and Interpretations.* France: Organisation for Economic Coopertaion and DEvelopment (OECD).

Stronge, J. H., & Tucker, P. D. (2020). *Handbook on Teacher Evaluation: Assessing and Improving Performance*. New York: Routledge. https://doi.org/10.4324/9781315505534

Wilujeng, I. (2018). *IPA Terintegrasi dan Pembelajarannya* (Vol. 1). Yogyakarta: UNY Press.