The construct validity and reliability of the lesson plan assessment instrument in primary schools

Lian Gafar Otaya [1]*, Badrun Kartowagiran [2], Heri Retnawati [2]

[1] IAIN Sultan Amai Gorontalo. Jalan Gelatik No.1 Kota Gorontalo, Gorontalo 96114, Indonesia
[2] Universitas Negeri Yogyakarta. Jalan Colombo No.1, Yogyakarta 55281, Indonesia
*Corresponding Author. E-mail: lianotaya82@iaingorontalo.ac.id

Abstract

This study aims to prove the construct validity of the lesson plan assessment instrument in primary schools. In addition, the purpose of this study is to estimate the reliability of lesson plan instruments in primary schools. This research uses a descriptive quantitative approach that is carried out on professional teacher education students at Universitas Negeri Gorontalo, Universitas Negeri Yogyakarta and Universitas Islam Negeri Makassar. The subjects of this study were 516 randomly selected students. Data collection is done through documentation of the results of assessments from the field supervisor of each student. The data analysis technique used is confirmatory factor analysis and composites score reliability. The results showed that the lesson plan assessment instrument was measured by 25 items spread over 4 indicators. All items in the lesson plan assessment instrument indicators are construct valid after being tested through confirmatory factor analysis. In addition, the lesson plan assessment instrument in this study was reliable and had a fairly high construct reliability coefficient of 0.92.

Keywords: construct validity, reliability, lesson plan assessment

Introduction

Teacher quality is seen as the most influential factor on the quality of education, because of the good and bad quality of education is very much determined by the teacher and becomes an important factor in preparing human resources (Bahcivan & Cobern, 2016; Gerritsen et al., 2016; LeCornu, 2016). Qualified teachers have a significant and substantive influence on the achievement of student learning outcomes. Teachers who work with high performance tend to have higher academic achievement (Steinberg & Garrett, 2016; Stronge, 2018). Therefore, it is important to improve the factors that affect teacher quality over time.

Since the past until the 4.0 industrial revolution that has been echoed today, improving teacher quality has remained a global concern of all countries in the world, especially in educational institutions that prepare prospective teachers not to be separated from criticism and consensus that questions teacher quality standards and how to improve them (Gareis & Grant, 2014; Good, 2008; Goodwin, 2010; Zhu et al., 2017). In addition, the demand to improve the quality of learning and high academic achievement every year has placed issues related to the professional development of teachers and teacher candidates to be an important concern by education observers, researchers and policy makers (Creemers et al., 2012; Rabadi-Raol, 2019). This shows the importance of improving the quality of teachers and prospective teachers in various countries, including in Indonesia, becoming a necessity for the development of its human resources.

One indicator of teacher quality is being able to prepare lesson plans well. Lesson plan is one important component in learning. This is important because it is a guide for teachers in learning activities at each meeting (Sanjaya, 2016). Therefore the success of the learning process also depends on good planning. Generally learning activities in a lesson plan are divided into three components , namely introduction, core, and closing (Fouryza et al., 2019). These three things are the core of the lesson plan.

Given the importance of the role of lesson plans in learning, then of course this should be one of the serious concerns for teachers and supervisors in planning learning for students. Ideally, each lesson plan should be evaluated as well as possible before applying it to learning. This of course requires appropriate measuring tools to control the quality of the lesson plan.

The problem so far is that information related to the quality of the lesson plan products for teachers has not been well documented. Assessments related to lesson plan quality tend to be based solely on the subjective opinions of supervisors. There might even be different perceptions among different supervisors even though they are evaluating the same lesson plan. In general, assessment is an important component of an education program (Herwin et al., 2019). Accurate assessment results will produce an objective policy. The absence of valid and reliable instruments that can be used to measure the quality of lesson plans that have been prepared by the teacher. Therefore, this research has developed a lesson plan assessment instrument and its validity and reliability will be proven. The purpose of this study is to prove the construct validity and reliability of the lesson plan assessment instrument.

Method

This research uses a descriptive quantitative approach. Location of research data collection was conducted at Universitas Negeri Gorontalo, Universitas Negeri Yogyakarta, and Universitas Islam Negeri Alauddin Makassar. The research was carried out in 2019. The subjects of this study were students of Teacher Professional Education in 2019.

The sample of this study were 516 randomly selected participants. This was done on the basis of the opinion of Comrey & Lee (1992) saying the sample size for factor analysis was 50 very bad, 100 bad, 200 fair, 300 good enough, 500 good, and 1,000 very good. Therefore the selection of 516 participants was good in conducting factor analysis.

The research data was collected through documentation of the results of the assessment conducted by the field supervisor. In addition, data analysis was performed with confirmatory factor analysis to prove the construct validity of the instrument. The general model of factor analysis measurement is built by the equation:

Indicator = λ Construct + Error

$x = \lambda\xi + \delta$

where:

x : Vector for indicator variables

ξ : Exogenous latent variable

λ : Loading factor

δ : Eror

(Nasir et al., 2015)

To test the significance of each indicator using the view of Retnawati (2016b) which states that the coefficient of the factor loading path will have meaningful if the magnitude is not less than 0.4. Verification of construct reliability is done using the composite score method or omega reliability. The formula to get the construct reliability coefficient (Retnawati, 2016a) is as follows.

$$\omega = \frac{\left(\sum_{i=1}^{i}\lambda_i\right)^2}{\left(\sum_{i=1}^{i}\lambda_i\right)^2 + \left(\sum_{i=1}^{i}1-\lambda_i{}^2\right)}$$

where

ω = Reliability coefficient

λ_i = Standardized loading factors

As a criterion used in reliability is 0.85 (Retnawati, 2016b). Goodness of Fit Model testing is done to test the suitability / identities between empirical data and theoretical models that are designed. The model is said to be fit if the emerald data is identical to the theoretical model. If p value $\geq$ 0.05 and RMSEA $\leq$ 0.08, the model can be explained as a fit model (Bentler, 1990; Hu & Bentler, 1999; McDonald & Marsh, 1990).

Results and Discussion

The lesson plan assessment instrument is measured through four indicators with 25 items distributed. The four indicators are indicators of the achievement of competency indicators and learning

outcomes (PIPKC) of 6 items (items 1-6), indicators of organizing material, methods, media and learning resources (PMMMS) of six items (items 7-12), organizing indicators learning process, assessment and evaluation (PPPEP) of six items (items 13-18), as well as indicators of the application of the principle of techno pedagogical content knowledge (PTPCK) of seven items (items19-25).

Indicators for the formulation of competency achievement indicators and learning outcomes (PIPKC) are measured by six items. These items consist of: completeness of writing identity, completeness of writing core competencies, suitability and clarity of the formulation of competency achievement indicators with basic competencies, clarity of the formulation of competency achievement indicators using verbs that can be measured and/or observed, suitability and clarity of learning objectives formulation with achievement indicators competence, as well as the completeness of the formulation of learning objectives meet the ABCD criteria (Audience, Behavior, Condition, Degree).

The indicators for organizing materials, methods, media and learning resources (PMMMS) were measured by six items. These items consist of: the suitability and clarity of the material with the learning objectives, the accuracy and completeness of the preparation of the material and teaching materials, the suitability of the choice of learning methods with basic competencies, the characteristics of the material and characteristics of the students, the suitability of the steps/syntax of learning with the learning strategies and the material taught, the suitability of the media and learning resources with learning objectives, materials, class conditions and the accuracy of the selection specifications, as well as the feasibility of the media and learning resources used.

The indicators for organizing the learning process, assessment and evaluation (PPPEP) were measured by six items. These items consist of: completeness and clarity of the steps of learning activities, clarity of learning scenarios describing active learning, clarity of learning scenarios reflecting scientific learning, appropriateness of assessment techniques with indicators of competency achievement and clarity of the scope of assessment, completeness of the assessment components used, and clarity of planned activities enrichment and/or remedial.

Indicators of applying the principle of techno pedagogical content knowledge (PTPCK) were measured by seven items. These items consist of: clarity of technological knowledge, clarity of pedagogical knowledge, clarity of content knowledge, clarity of techn ological pedagogical knowledge, clarity of technological content knowledge, clarity of pedagogical content knowledge, and clarity of technological pedagogical content knowledge.

Based on the results of the confirmatory factor analysis of the lesson plan assessment instrument, information was obtained that the Barlet coefficient with a p value of 0,000 and KMO of 0.9. If the results of the Barlet test show a p value less than 0.01 and a KMO coefficient of more than 0.05, then the adequacy of the sample for factor analysis has been fulfilled (Retnawati, 2016b). To prove the construct validity of each item, it is necessary to estimate the loading factor through confirmatory factor analysis. The following in Figure 1 presents the results of the confirmatory factor analysis.

Based on the information presented in Figure 1, it can be stated that as many as 25 items indicate their respective factor loads based on their indicators. Of the 25 items, it shows the entire factor load coefficient is more than 0.4. If this is confirmed by Retnawati (2016b)'s suggestion that the factor load coefficient will be meaningful if the magnitude is not less than 0.4. Thus, the results of the analysis show that all observable variables make a significant contribution to measuring the latent variables. This means that all 25 items in the lesson plan assessment instrument have been constructively valid.

Another thing that can be explained from the confirmatory factor analysis is the Goodness of Fit Model aspect. Based on the results of the analysis obtained information that the construct model that was designed had met the Goodness of Fit Model. This is indicated by the p-value coefficient of 0.052 which exceeds the 0.05 cut-of value and the RMSEA coefficient of 0.017 which is smaller than the 0.08 cut-of value. This was obtained after correlating errors between item 4 and item 15 and item 11 and item 19. Based on some of the analysis results that have been presented, it can be concluded that the lesson plan assessment instrument has been valid construct. Likewise the 25 items of the instrument have fulfilled the construct validity aspects.

The reliability of the lesson plan assessment instrument was analyzed using the reliability score composite approach using factor load coefficients for all observations of the latent variables. Following are the results of the reliability analysis of the lesson plan assessment instrument.
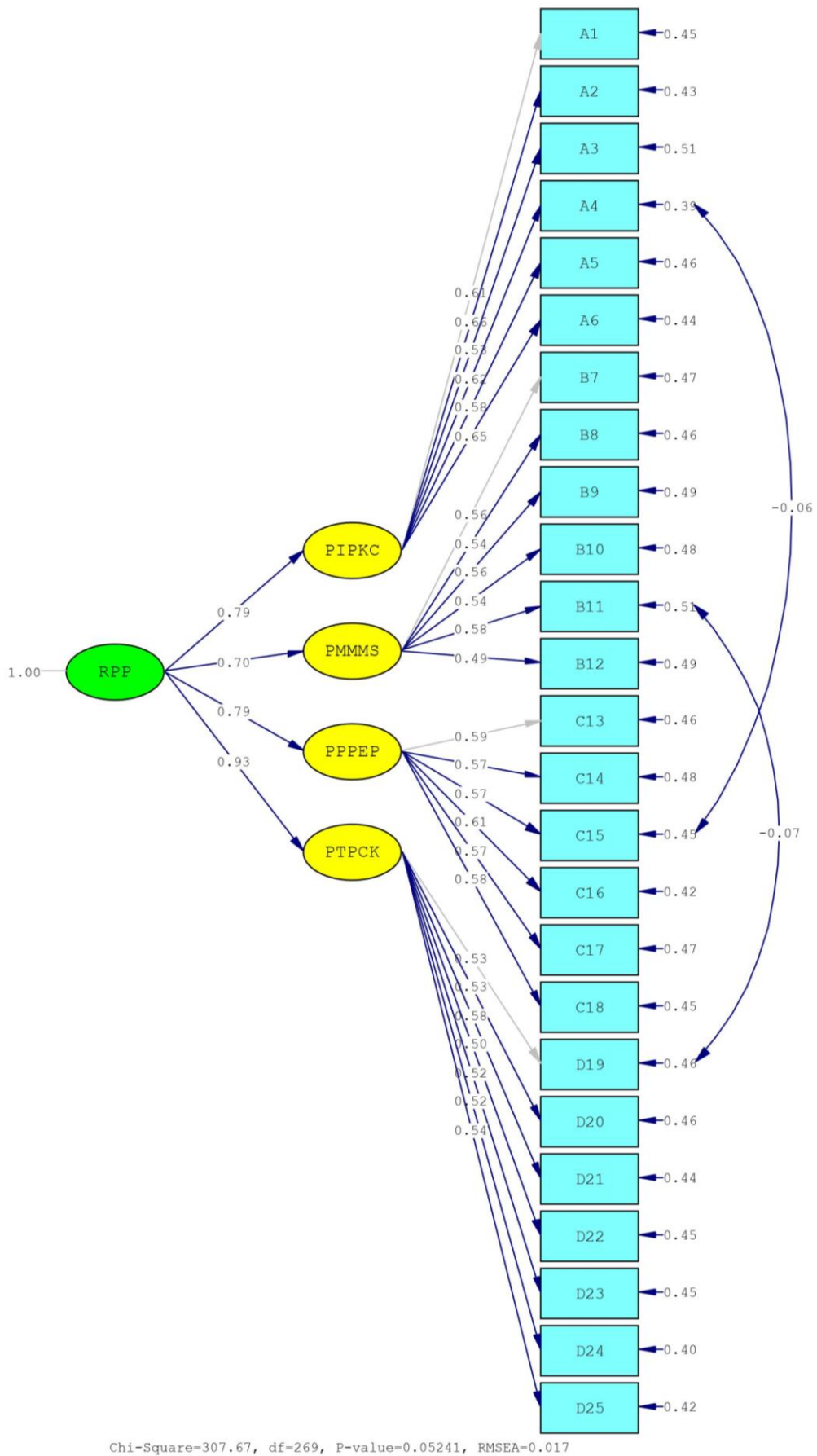
Chi-Square=307.67, df=269, P-value=0.05241, RMSEA=0.017

Figure 1. Confirmatory Factor Analysis Results

where
PIPKC   : Indicators of the achievement of competency indicators and learning outcomes
PMMMS   : Indicators of organizing material, methods, media and learning resources
PPPEP   : Organizing indicators learning process, assessment and evaluation
PTPCK   : Indicators of the application of the principle of techno pedagogical content knowledge

Table 1. Factor Loads of Observation Variables

Item	λ_i	λ_i^2	$1-\lambda_i^2$
A1	0,61	0,3721	0,6279
A2	0,66	0,4356	0,5644
A3	0,53	0,2809	0,7191
A4	0,62	0,3844	0,6156
A5	0,58	0,3364	0,6636
A6	0,65	0,4225	0,5775
B7	0,56	0,3136	0,6864
B8	0,54	0,2916	0,7084
B9	0,56	0,3136	0,6864
B10	0,54	0,2916	0,7084
B11	0,58	0,3364	0,6636
B12	0,49	0,2401	0,7599
C13	0,59	0,3481	0,6519
C14	0,57	0,3249	0,6751
C15	0,57	0,3249	0,6751
C16	0,61	0,3721	0,6279
C17	0,57	0,3249	0,6751
C18	0,58	0,3364	0,6636
D19	0,53	0,2809	0,7191
D20	0,53	0,2809	0,7191
D21	0,58	0,3364	0,6636
D22	0,50	0,2500	0,7500
D23	0,52	0,2704	0,7296
D24	0,52	0,2704	0,7296
D25	0,54	0,2916	0,7084
$\sum$	14,13		16,9693

Table 1 shows information about observational factor load values or partial scoring RPP assessment instrument items for lecturer evaluators in trial II. Based on the table, information obtained by factor load coefficient and squared factor load coefficient ω. This is used to calculate the reliability coefficient. In the table can be known $\sum_{i=1}^{i}\lambda_i$ of 14,13, and $\sum_{i=1}^{i}1-\lambda_i^2$ of 16,96. Through these coefficients, the reliability can be calculated as follows.

$$\omega = \frac{(14,13)^2}{(14,13)^2 + (16,96)}$$

$$\omega = \frac{199,65}{199,65 + 16,96}$$

$$\omega = 0,92$$

Based on the results of the reliability calculation obtained omega coefficient (ω) which shows the instrument reliability coefficient of 0.92. Mehrens & Lehmann (1973); Retnawati (2016b) states that although there is no agreement in general, it is widely accepted that the reliability coefficient must have a minimum criterion of 0.85. If based on these suggestions, it can be concluded that the lesson plan assessment instrument has been reliable.

The findings of this study indicate that the lesson plan assessment instrument has produced a construct with four indicators measured by 25 assessment items. Indicators for the formulation of competency achievement indicators and learning outcomes (PIPKC) are measured by six items. These items consist of: completeness of writing identity, completeness of core competency writing, conformity and

clarity of the formulation of competency achievement indicators with basic competencies, Clarity of formulation of competency achievement indicators using verbs that can be measured and/or observed, suitability and clarity of the formulation of learning objectives with achievement indicators competence, as well as the completeness of the formulation of learning objectives meet the ABCD criteria (Audience, Behavior, Condition, Degree).

The results of this study are relevant to the opinion of Archer and Hughes (2011) that explicitly teachers need to have the ability to choose learning material that must be taught and set criteria for success in learning and notify these criteria to students. This view shows that one of the abilities that must be possessed by the teacher is the ability to formulate competencies and determine learning outcomes. This means that in order to know the quality of learning planning undertaken by teachers, it is necessary to measure the ability of teachers to formulate competencies and determine learning outcomes.

The indicators for organizing materials, methods, media and learning resources (PMMMS) were measured by six items. These items consist of: the suitability and clarity of the material with the learning objectives, the accuracy and completeness of the preparation of the material and teaching materials, the suitability of the choice of learning methods with basic competencies, the characteristics of the material and characteristics of the students, the suitability of the steps / syntax of learning with the learning strategies and the material taught, the suitability of the media and learning resources with learning objectives, materials, class conditions and the accuracy of the selection specifications, as well as the feasibility of the media and learning resources used.

The results of the study were supported by Orazbayeva (2016) that ideally teachers must have the ability to effective teaching practices, namely implementing effective teaching practices by: (1) identifying, selecting and implementing various teaching strategies that can help students actively engage in learning; (2) conduct learning activities vary or are not monotonous; (3) identify the problems faced by students and help them solve them; (4) help students use learning resources; (5) provide lots of opportunities for students to ask questions, practice and interact with other students. These two dimensions are related to each other in supporting the success of learning carried out by the teacher.

This view shows the importance of organizing material, methods, media and learning resources. The ideal teacher is those who are able to organize the material well, choose learning methods correctly and determine learning resources that can support the achievement of learning objectives. Therefore, to find out the quality of learning planning undertaken by the teacher, it is necessary to measure his ability to organize the material, choose learning methods and determine the learning resources appropriately.

The indicators for organizing the learning process, assessment and evaluation (PPPEP) were measured by six items. These items consist of: completeness and clarity of the steps of learning activities, clarity of learning scenarios describing active learning, clarity of learning scenarios reflecting scientific learning, appropriateness of assessment techniques with indicators of competency achieve-ment and clarity of the scope of assessment, completeness of the assessment components used, and clarity of planned activities enrichment and/or remedial.

Theoretically the results of the study are supported by the views of Sanders et al (1990) stated that according to the American Federation of Teachers (AFT), the National Council on Measurement in Education (NCME), and the National Education Association (NEA) there are 7 (seven) standards for the scope of roles and responsibilities of teachers in assessing students namely: (1) skilled in choosing valuation methods; (2) skilled in developing valuation methods; (3) skilled in designing and interpreting assessment results; (4) skilled in using assessment results; (5) skilled in developing assessment procedures; (6) skilled in communicating or delivering assessment results; (7) skilled in recognizing inappropriate, unethical, illegal valuation methods and using valuation information.

In summary, the views of Sanders et al (1990) emphasizes the importance of organizing the process, assessment and evaluation of learning in a learning p lan. This is very important because organizing the right assessment will certainly give birth to the right results, information and decisions as well. If from the beginning the organization of the assessment was wrong, the decision related to the results of the assessment made by the teacher would also be wrong. This underlies the importance of organizing assessments and evaluations to be a measuring indicator in an assessment of the quality of learning plans developed by teachers.

Another indicator in the construct of the RPP assessment instrument is the application of the principle of techno pedagogical content knowledge (PTPCK) measured by seven items. These items

consist of: clarity of technological knowledge, clarity of pedagogical knowledge, clarity of content knowledge, clarity of technological pedagogical knowledge, clarity of technological content knowledge, clarity of pedagogical content knowledge, and clarity of technological pedagogical content knowledge.

The results of this study are relevant to the views of (Widiati & Hayati, 2015) that one of the principles that must be fulfilled in learning planning is TPACK (Technological Pedagogical and Content Knowledge) oriented, namely the integration of information technology, pedagogy, and content knowledge in the learning process. During the development of learning tools it is necessary to ensure that teachers implement TPACK, when selecting and determining strategies, approaches or models and learning media, must pay attention to the characteristics of students, materials, and learning objectives. Lecturers also need to ensure students use IT to improve learning effectiveness. Lecturers need to give examples of how to accommodate TPACK in learning. To implement the learning system it is necessary to pay attention to the principles of learning that are based on learning by doing, activeness, high-level thinking, accompaniment impacts, feedback mechanisms regularly, the use of information technology, contextual learning, the use of multistrategy and various learning resources, and oriented to TPACK . The integration of information technology, pedagogy, and content knowledge in the learning process has a very important role so that it must be stated in the planning of learning. Therefore to determine the quality of learning planning developed by teachers, it is necessary to measure indicators of Technological Pedagogical and Content Knowledge.

Based on some previous descriptions show that both empirically in the results of this study, as well as theoretically or concepts that have been formulated by relevant findings beforehand, it can be concluded that to conduct an assessment of learning planning the construct of the measurement indicators used include: the formulation of indicators of competency achievement and achievement learning (PIPKC), organizing materials, methods, media and learning resources (PMMMS), organizing processes, learning evaluation and evaluation (PPPEP), and applying the principles of techno pedagogical content knowledge (PTPCK).

The next aspect that becomes the unit of analysis in this study is the quality of the lesson plan instrument is reviewed based on the construct validity. The results of this study indicate that all items in the developed instrument have been construct constructively based on the construct validity criteria of an instrument. Conceptually Fernandes (1984); Nunnaly (1981); Retnawati (2016b) argues that construct validity indicates the extent to which the instrument reveals a certain theoretical ability or construct to be measured. If this view is related to the findings of this study it can be explained that the lesson plan assessment instrument that has been developed has been able to reveal the theoretical construct of the ability to be measured.

The findings of this study are also supported by the opinion of Furr and Bacharach (2013) which states that the construct validity refers to the extent to which the measurement results reflect the psychological construct being measured. This means that if an instrument has been proven to be construct valid, it has reflected the psychological construct to be measured. In relation to this research, the lesson plan assessment instrument has been able to reflect the ability of teachers or prospective teachers to be measured, namely the ability to plan learning.

Another aspect that becomes the unit of analysis in the lesson plan assessment instrument is the reliability of the instrument. Based on the results of this study indicate that the lesson plan assessment instrument has been reliable. This finding is supported by the opinion of Herwin and Mardapi (2017); Ziegler and Detje (2012) explaining that reliability reflects the overall consistency of the measurement even though it is given several times. Measurements that have high reliability are said to be reliable measurements. Reliability itself has other names such as trustworthiness, reliability, constancy, stability, consistency, and so on. However, the main idea contained in the concept of reliability is the extent to which the results of a measurement can be trusted.

The same thing is supported by Mehrens and Lehmann (1973); Retnawati (2016b) that a reliable measuring instrument will provide stable and consistent measurement results. This means that a measuring instrument is said to have a high reliability coefficient when used to measure the same thing at different times and the results are the same or close to the same. Based on the findings of this study it can be explained that the reliability coefficient of 0.92. This is evidence that the instrument that has been developed has shown a stability and consistency of measurement results.

Furthermore Retnawati (2016b) argues that reliability is also related to measurement error. High reliability indicates a small measurement error in obtaining measurement results. The greater the

reliability of an instrument, the smaller the measurement error and vice versa. If the opinion is related to the results of this study which shows the reliability coefficient of this instrument, it can be concluded that the reliability of the lesson plan assessment instrument is high enough so that the measurement error is very small. This shows a positive thing in terms of the development of the instrument and this is one that underlies researchers to conclude that the instrument is standard and is worthy of further use with good quality.

Conclusion

Based on the results of research and discussion it can be concluded that the lesson plan assessment instrument is measured through four indicators with 25 items distributed. The four indicators are indicators of the achievement of competency indicators and learning outcomes, indicators of organizing material, methods, media and learning resources, organizing indicators learning process, assessment and evaluation, and indicators of the application of the principle of techno pedagogical content knowledge. The lesson plan assessment instrument has been proven to be valid based on construct validity through confirmatory factor analysis. In addition this lesson plan assessment instrument has been reliable and has a fairly high reliability coefficient of 0.92.

Based on the conclusions of this study, it can be explained that the lesson plan is very important for educational programs. For this reason, the quality of the lesson plan must be good. It is recommended that before being implemented in learning activities, it is better if the lesson plan made by the teacher is evaluated first. The evaluation can be done using an assessment instrument that has proven its validity and reliability in this study. In addition, because the lesson plan assessment instrument is a positive one, it is recommended that this instrument be applied continuously in the field both for teachers and for prospective teacher students.

References

Archer, A., & Hughes, C. (2011). *Explicit instruction: efficient and effective teaching*. Guilford Publications.

Bahcivan, E., & Cobern, W. W. (2016). Investigating coherence among Turkish elementary science teachers' teaching belief systems, pedagogical content knowledge and practice. *Australian Journal of Teacher Education*, *41*(10), 63–68. https://doi.org/10.14221/ajte.2016v41n10.5

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, *107*, 238–246. https://doi.org/10.1037/0033-2909.107.2.238

Comrey, A. L., & Lee, H. B. (1992). *A First Course in Factor Analysis*. Lawrence Erlbaum Associates, Inc.

Creemers, B., Kyriakides, L., & Antoniou, P. (2012). *Teacher professional development for improving quality of teaching*. Springer Science & Business Media.

Fernandes, H. J. X. (1984). *Evaluation of educational program*. National Education Planning, Evaluating and Curriculum Development.

Fouryza, D., Amin, S. M., & Ekawati, R. (2019). Designing lesson plan of integer number operation based on fun and easy math (FEM) approach. *International Journal of Evaluation and Research in Education (IJERE)*, *8*(1), 103–109. https://doi.org/10.11591/ijere.v8.i1

Furr, R. M., & Bacharach, V. R. (2013). *Psychometrics: An introduction*. SAGE Publications.

Gareis, C. R., & Grant, L. W. (2014). The efficacy of training cooperating teachers. *Teaching and Teacher Education*, *39*, 77–88. https://doi.org/10.1016/j.tate.2013.12.007

Gerritsen, S., Plug, E., & Webbink, D. (2016). Teacher quality and student achievement: Evidence from a sample of Dutch Twins: Teacher quality and student achievement. *Journal of Applied Econometrics*, *32*(3). https://doi.org/10.1002/jae.2539

Good, T. L. (2008). *21st century education: A reference handbook*. Sage Publication Inc.

Goodwin, A. L. (2010). Curriculum as colonizer:(Asian) American Education in the current US context. *Teachers College Record*, *112*(12), 3102–3138.

Herwin, H., & Mardapi, D. (2017). An emotion assessment model for elementary school students.

Jurnal Penelitian Dan Evaluasi Pendidikan, *21*(1), 80–92. https://doi.org/10.21831/pep.v21i1.14504

Herwin, H., Tenriawaru, A., & Fane, A. (2019). Math elementary school exam analysis based on the Rasch model. *Jurnal Prima Edukasia*, *7*(2), 106–113.

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1–55. https://doi.org/10.1080/10705519909540118

LeCornu, R. (2016). Professional experience: learning from the past to build the future. *Asia-Pacific Journal of Teacher Education*, *44*(1), 80–101. https://doi.org/10.1080/1359866X.2015.1102200

McDonald, R. P., & Marsh, H. W. (1990). Choosing a multivariate model: Noncentrality and goodness of fit. *Psychological Bulletin*, *107*, 247–255. https://doi.org/10.1037/0033-2909.107.2.247

Mehrens, W. A., & Lehmann, I. J. (1973). *Measurement and evaluation in education and psychology*. Holt, Rinehart and Winston.

Nasir, N., Sappaile, B. I., & Arafah, K. (2015). Pendeteksian pelaksanaan supervisi akademik pengawas sekolah pada SMA Negeri di Kota Baubau melalui Analisis Faktor Konfirmatori (CFA). *Prosiding Konferensi Ilmiah Tahunan Himpunan Evaluasi Pendidikan*.

Nunnaly, J. C. (1981). *Psychometric theory*. Tata Mcgraw Hill Publishing Company Limited.

Orazbayeva, K. O. (2016). Professional competence of teachers in the age of globalization. *International Journal of Environmental and Science Education*, *11*(9), 2659–2672.

Rabadi-Raol, A. (2019). Quality of teacher education and learning: theory and practice. *Journal of Education for Teaching.*, 115–117.

Retnawati, H. (2016a). *Analisis kuantitatif instrumen penelitian: Panduan peneliti, mahasiswa dan psikometrian*. Parama Publishing.

Retnawati, H. (2016b). *Validitas reliabilitas dan karakteristik butir*. Parama Publishing.

Sanders, J. R., Nitko, A. J., Merwin, J. C., Trice, C., Dianda, M., & Schneider, J. (1990). *Standars for teacher competence in educational assessment of students*. NCME, American Association of Colleges for Teacher Education, American Federation of Teachers, and National Education Association.

Sanjaya, W. (2016). *Strategi pembelajaran berorientasi standar proses pendidikan*. Prenada Media. https://doi.org/2008

Steinberg, M. P., & Garrett, R. (2016). Classroom composition and measured teacher performance: What do teacher observation scores really measure? *Educational Evaluation and Policy Analysis*, *38*(2), 1–25. https://doi.org/10.3102/0162373715616249

Stronge, J. H. (2018). *Qualities of effective teachers*. ASCD.

Widiati, U., & Hayati, N. (2015). *Teacher professional education in Indonesia and ASEAN 2015: Lessons learned from English language teacher education programs* (R. Stroupe & K. Kimura (eds.)). ASEAN integration and the role of English language teaching.

Zhu, X., Goodwin, A. L., & Zhang, H. (2017). *Quality of teacher education and learning: Theory and practice*. Springer Nature.

Ziegler, J., & Detje, F. (2012). Aplication of empirical methodology to evaluate information fusion aproaches. *International Journal Methodology*, *20*, 327–337.