The Implementation Of Generalisability Theory On Physics Teachers' Competency Assessment Instruments Development

Syahrul Ramadhan, Subair Ahmad Nasran, Heru Budi Utomo, Faridl Musyadad, Suryadi Ishak

Abstract: This research was conducted to make a standard instrument in assessing physics teacher competencies. The structure of this assessment instrument is four types of competencies that must be possessed by physics teacher (social, pedagogic, professional, personality). Determination of coefficients the reliability of the assessment instrument is conducting by using a Genova computer program package based on the generalizability theory developed by Crick and Brennan. The research subjects were 30 physics teachers in the district of Bima NTB and involved four experts as assessors. The G coefficient of this instrument is 0.738. Furthermore, the analysis of the D Study concludes that to reach an understanding and agreement that meets the level of observation that is acceptable for facet broader, namely > 0.70.

1. INTRODUCTION

Education is the primary key to developing a country. Education creates competent human resources capable of competing in science and technology development. The development of science and technology forces all nations to improve education quality. The teacher is the main component to determine the progress of education. Therefore, teacher quality is needed to maintain and improve along with the times. The learning activity determines teacher quality. Although the government provides a standard policy for the education process as an effort to improving education quality, the teacher will implement and determine learning quality that will impact education quality in general. Therefore, an effort to standardise education in Bima Regency, it is necessary to conduct knowledge to find out the standardisation of teacher competencies, especially physics teachers. To assess these competencies, an instrument is needed to measure the competence of physics teacher authentically. Many instruments development as teacher competency parameter have been done, such as Rahmat, Riandi, Solihat, WB, Zaputra and Ferazona [1] who developed a high school Biology Teacher competency instrument based on analysis of the suitability of learning in the classroom with the demands of Basic Competence (KD). Another study was conducted by Manutede, Susiloningsih and Ridlo [2] about developing a valid and reliable self-assessment instrument for pedagogical competence and professional competence that was implemented in teachers in junior high schools in Salatiga City. These studies can be able to produce valid and reliable instruments. However, the instrument does not reflect authentic and specific teacher's abilities. This research was conducted to make a standard instrument in assessing physics teacher competencies. Developing instruments to assess teacher competence physics begins with producing a measurable structure. The structure of this assessment instrument is four types of competencies that must be possessed by physics teacher (social, pedagogic, professional, personality). To test the measured structure, the researcher used the generalisability theory. The use of this theory is considered good because it is able to reveal the relationship between each variable and error source comprehensively.

2. SYSTEMATIC REVIEW

2.1 Teacher Competence

Law Number 20 of 2003 [3] about National Education System states that educators are professionals. Therefore, the teacher as a professional educator has a strategic function, role, and position. The teacher as professional has the vision to realise the learning implementation based on the professionalism principles to fulfil the same rights for everyone in obtaining a quality education. Law Number 14 of 2005 about Teachers and Lecturers define that teachers must have academic qualifications, competencies, educator certificates, physical and spiritual health, and fulfil other qualifications required by the higher education unit where they are assigned and can actualise the national education goals. Government Regulation Number 74 of 2008 article 3 paragraph 1 about teachers. Teacher competencies as stated in paragraph (1) include pedagogical competencies, personality competencies, social competencies, and professional competencies conducted through professional education. A pedagogical competence is an ability to manage the learning of learners, including understanding the learners, the design and implementation of learning, evaluation of learning outcomes, and the development of students to actualise their potential. This competency consists of 4 aspects assessed, such as knowing the student's characteristics, mastering learning theory and the principles of learning, understanding and developing students potential and communication with students. The personality competence is a steady, stable, mature, wise and dignified, become role models for students, and great personality of an educator. In personality competence, several aspects are assessed, such as mature, noble character, becoming a role model for students and society, and sustainable self-development. The social competence is educators ability to communicate and interact effectively with students, fellow teachers, staff, parents/guardians of students, and the community. Social competence is assessed by, a) Associating effectively with students, fellow educators, education personnel, parents/guardians of students. b) Acting based on the norms of religion, law, social, and national culture of Indonesia c) Showing maturity and role model d) Work ethic, high responsibility, feeling proud to be a teacher. The professional competence is the capacity of educators in the mastery of

learning materials is broad and deep that enable it to guide learners to gain the competency. The aspects of professional competence assessed including a) Concepts, structures, scientific methods that are coherent with the material, b)Teaching materials that exist in the school curriculum, c) Relationship of concepts between related lessons and d) Scientific concepts implementation in daily life.

2.2 Generalizabilitas Theory

Determination of coefficients the reliability of the assessment instrument done by using a package of Genova computer programs based on the generalise ability theory developed by Crick and Brennan in 1983 called A Generalized Analysis of Variance System [5]. In this theory, there is a G (generalised study) and D (decision study). In G-study estimates were made several component variances. The number of components is determined by the model used. Results of G-study is used in D-study. D-study emphasises the estimation, use, and interpretation of component variances to make decisions, with measurement procedures the good one. The important thing in the D-study is the specification of the generalisation of the universe, the universe of enactment D-study generalisation with a particular measurement procedure [5-7]. This research using the GENOVA component with the variation are person, rater, item, person and rater interaction, and error [5]. G study uses nested design, and so does D-study. This study uses one facet px (i: r) of G-study which is nested to estimate component variances, error variances, generalizability and phi coefficients for one-facet, nested, i: r D-study. The component variation that blend in nesting design (p, r: i, e) is the number of variances component in nested G-study is written as follows.

$$\sigma^2_{p,r:i,e} = \sigma^2_p + \sigma^2_{r:i,e}$$

Description : p = person, r = teacher / rater, i = item, r: i = rater nesting on items, e = error After the component variation is conducted, including the error variety, true variety can be estimated. Furthermore, it can be estimated that the magnitude of the reliability index of the measurement results, namely the actual ratio variety to the overall component variety. Estimated variation of each component and the magnitude of the reliability index measured by the instrument developed by the researcher using the GENOVA program package. The design used for G-study is px (i: r), which is a nesting item on the rater, the expert in assessing the results of the competency of the teacher interaction with the teacher who lodges in the item. The way the assessor (rater) assesses the competency of the teacher (p) depends on the opinion the assessor of the item being assessed so that the rater is said to nest on the item. The design of px (r: i) is based on a variant analysis random effects have a main effect: p, r, r: i and the interaction effect is pi, pr is nested on i. So there are person variant, rater variant, and rater variant nesting on i for the main effect, while the interaction effect is the variance person item, the rater variant that is lodged on the item . The amount of variance r nested in i can be written as follows.

$$\sigma^2 (r : i) = \sigma^2 (r , ri) = \sigma^2 (r) + \sigma^2 (ri).$$

The magnitude of the reliability coefficient of the assessment instrument is:

$$E\rho^2 = \frac{\sigma^2 (p)}{\sigma^2 (p) + \sigma^2 (\delta)}$$

$E\rho^2$ is the expected value of the instrument reliability coefficient,

$\sigma^2 (p)$ is the person variance,

$\sigma^2 (\delta)$ is error variance.

3. RESEARCH METHODS

This research is development research which uses a quantitative and qualitative approach. Development research used to conduct standard instruments in assessing physics teacher competencies. The research subjects were 30 physics teachers in the district of Bima NTB and involved four experts as assessors. These experts are master in the field of sociology (representing social competencies), master in research and educational evaluation (representing pedagogical competencies), physicists (representing professional competencies) and psychological master (representing personality competencies). Determination of coefficients the reliability of the assessment instrument is conducting by using a Genova computer program package based on the generalizability theory developed by Crick and Brennan in 1983 called A Generalized Analysis of Variance System [5].

4. ANALYSIS AND DISCUSSION

4.1 Model Design

This research using the GENOVA component the variant are person, rater, item, person and rater interaction, and error. G study uses nested design, and so does D-study. This study uses one facet p x (i: r) G-study that is nested to estimate component variant, error variant, generalizability and phi coefficients for one-facet, nested, i: r D-study. The design used for G-study is px (i: r), which is a nesting item on the rater in assessing physics teacher competencies. The way of rater assesses teacher competency (p) depends on the assessor view on the item being assessed (competency), so the rater is lodged on the item. The design of px (r: i) is based on a variant analysis random effects have a main effect: p, r, r: i and the interaction effect is pi, pr is nested on i. So there are person variant, rater variant, and rater variant nesting on i for the main effect, while the interaction effect is the variance person item, the rater variant that is lodged on the item. At this initial draft, the assessment instrument is tested to thirty teachers (P) and composed of four experts as rater (R) or assessor of teacher competency. Each rater assesses different competencies, and each competency consists of four items/aspects (I).

4.2 Genova Analysis Results

Summary of G study analysis from trial data

TABLE 1

Summary of G study analysis from trial data

	VARIANCE COMPONENTS IN TERMS OF G STUDY UNIVERSE (OF ADMISSIBLE OBSERVATIONS) SIZES					VARIANCE COMPONENTS IN TERMS OF D STUDY UNIVERSE (OF GENERALIZATION) SIZES				
EFFECT	VARIANCE COMPONENTS FOR SINGLE OBSERVATIONS	FINITE UNIVERSE COR-RECTIONS	D STUDY SAMPLING FRE-QUENCIES	VARIANCE COMPONENTS FOR MEAN SCORES ESTIMATES	STANDARD ERRORS	VARIANCE COMPONENTS FOR SINGLE OBSERVATIONS	FINITE UNIVERSE COR-RECTIONS	D STUDY SAMPLING FRE-QUENCIES	VARIANCE COMPONENTS FOR MEAN SCORES ESTIMATES	STANDARD ERRORS
P	0.00000E+00	1.0000	1	0.00000E+00	0.88209	2.57247	1.0000	1	2.57247	0.71417
R	0.28895	1.0000	4	0.07224	0.12629	0.28895QFM0000E+00		4	-------	-------
I:R	0.12850	1.0000	16	0.00803	0.01470	0.12850	1.0000	16	0.00803	0.01470
PR	10.28987	1.0000	4	2.57247	0.52682	10.28987	0.0000E+00	4	-------	-------
PI:R	14.58678	1.0000	16	0.91167	0.06892	14.58678	1.0000	16	0.91167	0.06892

QFM = QUADRATIC FORM

Condition above is related to object measurement for physics teacher competence, which has a major influence on variant error measurement is person cross items that depend on rater (Px I: R). Source variance depends on rater (I: R) person cross items are component the most dominant variant. This is might because the teacher who becomes a person and expert who is the rater or evaluator know the model and structure tool assessment development.

4.3 Data Analysis Results G Study (Coefficient G)

TABLE 2

Data Analysis Results G Study (Coefficient G)

	VARIANCE	STANDARD DEVIATION	STANDARD ERROR OF VARIANCE	
UNIVERSE SCORE	2.57247	1.60389	0.71417	
EXPECTED OBSERVED SCORE	3.48414	1.86659	0.71084	
LOWER CASE DELTA	0.91167	0.95482	0.06892	GENERALIZABILITY COEFFICIENT = 0.73834 (2.82170)
UPPER CASE DELTA	0.91970	0.95901	0.06818	PHI = 0.73664 (2.79706)
MEAN	0.12417	0.35238		

NOTE: SIGNAL/NOISE RATIOS ARE IN PARENTHESES

GENOVA VERSION 3.1 PAGE

G study results for knowing the advantage of using teacher competency assessment instrument. G coefficient of components judgment it shows that in a manner whole instrument model. Development assessment of physics teacher competencies can be accepted for used do judgment on the facet which is larger or with said other has been meeting for interests facet related measurements with object measurement (universe of admissible observations) on assessment of physics teacher competencies named index G coefficient of 0.73834.

4.4 Data Analysis of Study D Results

The purpose of analysis D study is answering question design that the study need to be selected and the number of item component judgment should be covered to measure and rate teacher competence to show the advantage for more facets large. By looking at every stage design D study on composition big sample to obtained information of G coefficient and also obtained information of the improvement in the index of the advantages on coefficient G after one item component judgment involved for measure or rate. To answer these questions and the purpose implied inside, the analysis on every results D can be used. The description explained the results analysis D of this study.

TABLE 3
Summary results analysis D-Study

D STUDY #3 -- P X (I:R) DESIGN -- I RANDOM, R FIXED

SUMMARY OF D STUDY RESULTS FOR SET OF CONTROL CARDS NO. 003

D STUDY DESIGN NO	SAMPLE SIZES INDEX= UNIV.= $P INF.	R 4	I INF.	VARIANCES UNIVERSE SCORE	EXPECTED OBSERVED SCORE	LOWER CASE DELTA	UPPER CASE DELTA	MEAN	GEN. COEF.	PHI
003-001	30	4	1	2.57247	6.21916	3.64670	3.67882	0.23943	0.41364	0.41151
003-002	30	4	2	2.57247	4.39582	1.82335	1.83941	0.16259	0.58521	0.58308
003-003	30	4	3	2.57247	3.78803	1.21557	1.22627	0.13698	0.67910	0.67719
003-004	30	4	4	2.57247	3.48414	0.91167	0.91970	0.12417	0.73834	0.73664

GENOVA VERSION 3.1 PAGE 39

CONTROL CARD INPUT LISTING

Genova for a test tries judgment competence physics teacher can serve on Pictures and Tables following.

TABLE 4
Estimated Coefficient of Generalizability Competency Assessment of Physics Teachers
P x (I: R) Design - I Random, R Fixed

D STUDY DESIGN NO	SAMPLE SIZE			GENERALIZABILITY		Selisih Koefisien Genova
	$ P INF.	R INF	I INF.	COEF.	PHI	
003-001	30	4	1	0,41364	0,41151	0,00213
003-002	30	4	2	0,58521	0,58308	0,00213
003-003	30	4	3	0,67910	0,67719	0,00191
003-004	30	4	4	0,73834	0,73664	0,0017

The table above gives an overview of changes in the Generalizability coefficient for various compositions of sample sizes P, R, and I. For the physics teacher competency assessment component P x (I: R) Design - Random, R Fixed (D study design number 00 3 -001 with P = 3 0, R = 4 and I = 1). Then the level or coefficient of understanding and agreement (reliability in the G coefficient) is 0, 41364. That means the assessor has a level of understanding and agreement on the use of the construct of the assessment instrument used by 41 %. If the assessor uses two indicators (design D study)number 00 3 -002, with P = 3 0, R = 4 and I = 2) namely indicators / competencies 1 and 2, then the level or coefficient of agreement and agreement with total 0.58521 ; and so on for the draft 00 3 -003 coefficient of 0.67910. Based on this fact, it can be said that to achieve understanding and agreement that meets the level of observation which accepted for broader facets, namely > 0.70, the assessor must use indicators 1, 2, 3 and 4 because in this context if 4 indicators are used they will the understanding coefficient and agreement reached 0.73834 or 74 % (> 70%) . The use of generalisability theory in assessing the reliability of the development of the instrument has been widely used. It includes research conducted by Retnowati [8], which examines the development of assessment instruments to measure the learning outcomes of children's painting. Another study was conducted by Guntur, Sukadiyanto and Mardapi [9] who developed a valid and reliable assessment instrument to measure the physical education learning outcomes of sports and the health of high school students in volleyball games. The researchers agreed that the use of generalizability theories emphasizes the estimation, use, and interpretation of component variances to make decisions. Based on these reasons, the use of theory G is in accordance with good measurement procedure.

5. CONCLUSION

Based on the data and analysis above, the characteristics of physics teacher competency assessment instruments include the generalizeability theory (Theory G), and Decision theory (Theory D) can be known. The G coefficient of this instrument is 0.738. Furthermore, the analysis of the D Study concludes that to reach an understanding and agreement that meets the level of observation that is acceptable for facet broader, namely > 0.70, the assessor must use indicators 1 , 2, 3 and 4. Random designation P x (I: R) - I Random, R Fixed (P = 3 0, R = 4 and I = 4) has fulfilled the minimum criteria required are 0.70. Therefore, to get the results of an assessment of authentic physics teacher competencies, the instrument developed can be used , by involving four competency indicators. In the future, it is necessary to conduct in-depth research and analysis and a larger sample to provide a comprehensive information about the quality of developed instruments.

REFERENCES

[1] Rahmat, A., Riandi, R., Solihat, R., WB, W.W., Zaputra, R., and Ferazona, S.: 'Peta kompetensi guru biologi di sma kota bandung berdasarkan analisis kesesuaian proses pembelajaran di kelas dengan tuntutan kompetensi dasar', Jurnal Pengajaran MIPA, 2014, 19, (2), pp. 179-187

[2] Manutede, Y.Z., Susiloningsih, E., and Ridlo, S.: 'Pengembangan Instrumen Kompetensi Pedagogis Guru SMP dalam Kurikulum 2013 Menurut Persepsi Guru di Kota Salatiga', Journal of Educational Research and Evaluation, 2015, 4, (2)

[3] Indonesia, P.R.: 'Undang-undang Republik Indonesia nomor 20 tahun 2003 tentang sistem pendidikan nasional', Jakarta: Pemerintah Republik Indonesia, 2003

[4] Nomor, P.P.R.I.: 'tahun 2008 tentang Guru', in Editor (Ed.)^(Eds.): 'Book tahun 2008 tentang Guru' (edn.), pp.

[5] Crick, J., and Brennan, R.: 'GENOVA: A generalized analysis of variance system', Computer software and manual]. Iowa City, IA: University of Iowa.(Available on http://www. education. uiowa. edu/centers/casma/), 1983

[6] Brennan, R.: 'Elements of Generalizability Theory (Iowa City, IA', in Editor (Ed.)^(Eds.): 'Book Elements of Generalizability Theory (Iowa City, IA' (ACT Publications) p, 1983, edn.), pp.

[7] Brennan, R.L.: 'Generalizability theory', Educational Measurement: Issues and Practice, 1992, 11, (4), pp. 27-34

[8] Retnowati, T.H.: 'Pengembangan instrumen penilaian karya seni lukis anak di sekolah dasar', Jurnal Penelitian dan Evaluasi Pendidikan, 2009, 13, (1)

[9] Guntur, G., Sukadiyanto, S., and Mardapi, D.: 'Pengembangan asesmen hasil belajar penjasorkes siswa sma pada permainan bolavoli', Jurnal Penelitian dan Evaluasi Pendidikan, 2014, 18, (1), pp. 13-29