



Development of Peer Assessment Social Behaviour Instrument on Learning Mathematics in Junior High School

Maria Gerrin W^{1✉}, Kartono², Saiful Ridlo²

¹ SMP Bhakti Tunas Harapan, Magelang, Indonesia

² Universitas Negeri Semarang

Info Artikel

Sejarah Artikel:
Diterima 12 May 2018
Disetujui 31 July 2018
Dipublikasikan 24
August 2018

Keywords:
Instrument
development, patient
satisfaction, service
quality

Abstract

Social behaviour is one of the aspects assessed in the process of applying the Curriculum 2013. Social behaviour includes aspects of honest, disciplined, courtesy, confident, caring, and responsibility. This study aims to develop an instrument of social behaviour through peer assessment on mathematics learning in junior high school that are valid and reliable. The research method used is development research with modified Sugiyono model. Content validity through expert judgment is tested using V index and its reliability with ICC. Small-scale trials were tested with pearson correlation and alpha cronbach for reliability. Large-scale trials were tested using Confirmatory Factor Analysis 2nd Order, and its reliability with the Construct Reliability formula. The results of the validation of experts obtained 35 items valid and reliable. A small-scale trial with 52 respondents, 27 valid grains with coefficient > 0.3 and reliability were 0.805. In large scale trials, 25 items have construct validity with estimated loading factor (λ) > 0.30. The test result using SEM meets Godness of Fit ie RMSEA show value 0,017; GFI shows a value of 0.87, AGFI shows a value of 0.84 and RMSR shows a value of 0.041. Instrument expressed practical with the questionnaire score > 21. The results showed that the instrument of social attitudes through peer assessment developed valid and reliable.

© 2018 Universitas Negeri Semarang

✉ Alamat korespondensi:
Jl Jendral Sudirman No.68A, Magelang, Indonesia
E-mail: gerinnmaria@gmail.com

INTRODUCTION

The curriculum changes from Kurikulum Tingkat Satuan Pendidikan 2006 (KTSP 2006) to the Curriculum 2013 provide some important impacts in the teaching and learning process. One of the fundamental changes is an update to the Graduate Competency Standards. The Competency Standards of graduates in Curriculum 2013 includes three domains: behaviours, knowledge, and skills formulated in the four Core Competencies (KI). Core competence is the basis for the development of Basic Competence. The four Core Competencies are spiritual behaviours (KI-1), social behaviours (KI-2), knowledge (KI-3), and skills (KI-4). The Graduate Competency Standards in the Curriculum 2013 will be the basis for developing Basic Competencies in each subject and at each grade level.

Assessment of learning outcomes by educational unit aims to assess the achievement of Graduate Competency Standards for all subjects by considering the results of educator's appraisal of learners. Assessment of learning outcomes in the Curriculum 2013 is done by educators and educational units through the stages of assessing the syllabus as a reference for assessment planning, making the instrument grid and the determination of assessment criteria, the implementation of the assessment in the learning process, analyzing the assessment results and giving follow-up on the assessment made by educator, compile report of assessment result in the form of description of competency achievement and description of behaviour (Setiadi, 2016).

The 2013 curriculum develops two modes of learning process, namely direct learning process and indirect learning process. The process of direct learning is the educational process whereby learners develop knowledge, thinking skills, and psychomotor skills through interaction. Indirect learning is an educational process that occurs during a direct learning process but is not designed in a specific activity.

Both direct and indirect learning take place in an integrated and indivisible manner. In this indirect learning will be related to the development of values and behaviour. The development of this behaviour is carried out by all subjects including mathematics, so that the process in learning aims to develop morale and behavior related to the behaviour contained in core competence and basic competence (Asmarawati, et al, 2016).

Based on this the teacher must prepare various assessment instruments that will be used to assess the overall competence of behaviour, knowledge, and skill competencies. behaviours competence assessments conducted by teachers can be done through various ways that is through observation, self-assessment, peer assessment and journal notes. Instruments used for observation, self-assessment, and assessment among learners are checklists or rating scales with rubrics, whereas in journals are educators' records. Instruments of self-assessment and assessment among friends can be done at least once in a semester. Assessment of social behaviour that have been done by teachers have not used a measurable instrument that is a valid and reliable instrument. Social behaviour in the instruments are not developed according to the indicators of competence. Therefore, to be able to measure social behaviours it is necessary assessment instrument that can measure indicators of social behaviours in accordance with Basic Competence of learning. Measurement of social behaviours using rating instruments in the form of rating scale. Assessment of social behaviours conducted by involving students through peer assessment.

One method of assessment that has been effectively tested is peer assessment and self assessment. The application of peer assessment is usually done in group activities so as to enable students to observe each other (Zevenbergen, 2001). Students become convinced that peer assessment is useful, motivating, interesting and not boring. Integrating peer assessment into evaluation procedures not only encourages

learners and teachers to view assessment as a shared responsibility, but can also be applied to transform teacher-centered learning into more student-centered (Azarnoosh, 2013).

The development of instruments is an activity to develop existing instruments to be better than ever valid and reliable. Development of this instrument becomes important in determining the quality or information of a public service, especially in puskesmas on the quality of services provided. Development of the instrument must be through the stages of good development in order to obtain a quality instrument (Rusilowati, 2013). Instrument is a very important tool in a study, because with the instrument a data and information will be obtained. Therefore, to obtain good quality data, then the instruments used must be qualified that meet the validity and reliability requirements (Sugiyono, 2010).

Teachers have the responsibility to reflect and evaluate student learning outcomes by measuring the extent of their competencies, both in the cognitive and affective domains (Gulikers, 2009). Implementation of Curriculum 2013 which requires teachers to be able to assess students' skills that must cover the four core competencies make teachers difficult to be able to assess the learners (Dinar, 2016).

The social behaviour assessment instruments used by teachers in Magelang all use daily journals and some teachers use an observation sheet whose indicators are less appropriate to the aspect of social behaviour assessment to be achieved. Teachers have difficulty developing behaviour assessment indicators. In addition, the instruments used are not yet standardized because they have not passed the validity and reliability stages of the instrument. As a result, the data generated by the instrument that has not been valid and reliable can be said to be less good data that impact on the clarity and accuracy of information provided on the attainment of student behaviours. Measurement activities certainly can not be separated from the data

collection instrument that will be used as a basis for collecting the necessary data. The quality of the data also depends on the quality of the instruments used to collect the data so that the validity and reliability requirements must be met (Widoyoko, 2016).

Khuriyah (2003) form of instruments in the preparation of affective domain evaluation can not be separated from existing materials. However, the difference is that the form of the statement is more directed to the cases that are often experienced by the respondents everyday by using a language that is easily understood and digested by them. Mansyur (2011) assessment in the learning of mathematics is an integral part of the mathematics education process. Efforts to improve the quality of mathematics education can be pursued through improving the quality of learning and the quality of the assessment system. The quality of this learning can be seen from the assessment results. Conversely, a good assessment system, will encourage teachers to determine the right strategy and motivate students to learn better. One form of assessment that is expected for this is formative assessment.

This study aims to develop a reliable and reliable social behaviours instrument through peer assessment. The instrument developed in the form of an observation sheet with a form of social behaviours scale used through peer assessment. The social behaviours to be measured in the instruments are honest, disciplined, courteous, confident, caring, and responsible.

METHOD

This research was conducted at SMP Bhakti Tunas Harapan and SMP Negeri 8 Magelang during July 2018. The subjects were 200 students of class VIII. Development of instrument used refers to Sugiyono's development model consisting of 10 langkah but only used 9 steps, namely (1) Potential and problem, (2) Data collection, (3) Product design, (4) Design validation, (5) Repair design, (6)

Product test, (7) Revision of product, (8) Test of use, (9) Revision of product. The instruments developed include 6 aspects of social behaviours that are honest, disciplined, courteous, confident, caring, and responsibility consisting of 35 items of statement before testing validity and reliability.

Test method of content validity using expert valuation method and its reliability using Intraclass Correlation Coefficient (ICC). Validity and reliability on small-scale trials using Pearson correlation and reliability using Alpha Cronbach reliability test and construct validity on large-scale trials using Confirmatory Factor Analysis (CFA) and Construct Reliability analysis to measure reliability.

RESULTS AND DISCUSSION

The results of this study consist of the results of content validity test, construct validity and reliability of social behaviours assessment instruments through peer assessment.

Content Validity, Small Scale Trial, Large Scale Trial

Content Validity

Before field testing, the instrument must first be tested for the validity of the content based on the expert assessor in order to assess the contents of the developed instrument. Assessment involves 3 experts who are competent in the field of instrument development and measurement of social behaviours. Assessment is done by the expert by giving a score on the validation sheet that has been provided. The scores obtained were then analyzed using Aiken V formula. The results of Aiken V analysis for social behaviours assessment instruments can be seen in Table 1.

Table 1. Assessment of Instrument Validator of Social Behaviours Based on Item

No Item	V indeks	Criteria	No Item	V indeks	Criteria
1	0,889	Valid	19	0,889	Valid
2	0,667	Valid	20	0,889	Valid
3	1,000	Valid	21	1,000	Valid
4	0,778	Valid	22	0,889	Valid
5	1,000	Valid	23	0,333	Valid
6	1,000	Valid	24	0,889	Valid
7	1,000	Valid	25	0,556	Valid
8	0,889	Valid	26	0,444	Valid
9	0,333	Valid	27	1,000	Valid
10	0,444	Valid	28	0,444	Valid
11	1,000	Valid	29	0,444	Valid
12	0,778	Valid	30	1,000	Valid
13	0,889	Valid	31	0,778	Valid
14	0,667	Valid	32	1,000	Valid
15	1,000	Valid	33	1,000	Valid
16	0,667	Valid	34	1,000	Valid
17	1,000	Valid	35	0,889	Valid
18	1,000	Valid			

Based on the assessment table of experts all instrument items obtained V Index. If the item has a correlation index < 0.3 then it is declared invalid, in accordance with the theory expressed by Widoyoko (2016: 170) stating that the valid item is having correlation index > 0.3 , while the invalid has correlation index < 0.3 . Reliability used by Intraclass Correlation Coefficient (ICC) analysis, instrument reliability value according to experts is 0,708.

Small Scale Trial

Validated instruments were then conducted on a small scale trial conducted on 52 students. The validity of the grains was analyzed by the correlation correlation parson formula. The result of analysis of social behaviours assessment instrument on small-scale trial was 27 items of correlation instrument aitem total score 0.3. There are 8 items of invalid instrument or a correlation score of a total score

of less than 0.3 in 3, 6, 12, 15, 19, 22, 29, and 32. Thus the item is declared invalid and must be disqualified. Reliability analysis of social behaviours assessment instruments using Alpha Cronbach reliability test. Alpha Cronbach reliability measurement results obtained value of instrument reliability coefficient of 0.805. This small-scale trial obtained 27 valid grains and can be used for large-scale trials.

Large Scale Trial

Instruments that have been revised in small-scale trials are then tested on a large scale in 148 students. This is done to test the validity of instrument constructs. Analysis of construct validity using Confirmatory Factor Analysis (CFA) second order with help of LISREL program. The value of the validity of the grains is seen from the SLF (standardized factor loading) value which can be seen in the path diagram model of Figure 1.

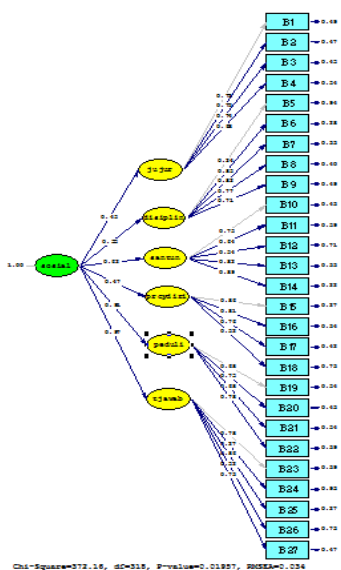


Figure 1. Output Lisrel Path Diagram Factor Value Loading Instrument

Based on Figure 1 there are 2 items of statement that are below 0.3 so that the grains will be discarded to meet the criteria of goodness of fit that have been specified Model path

analysis formed from the collapse of social behaviours can be seen in Figure 1.

After item 5 and item 24 deleted, the output of lisrel factor loading can be seen in Figure 2.

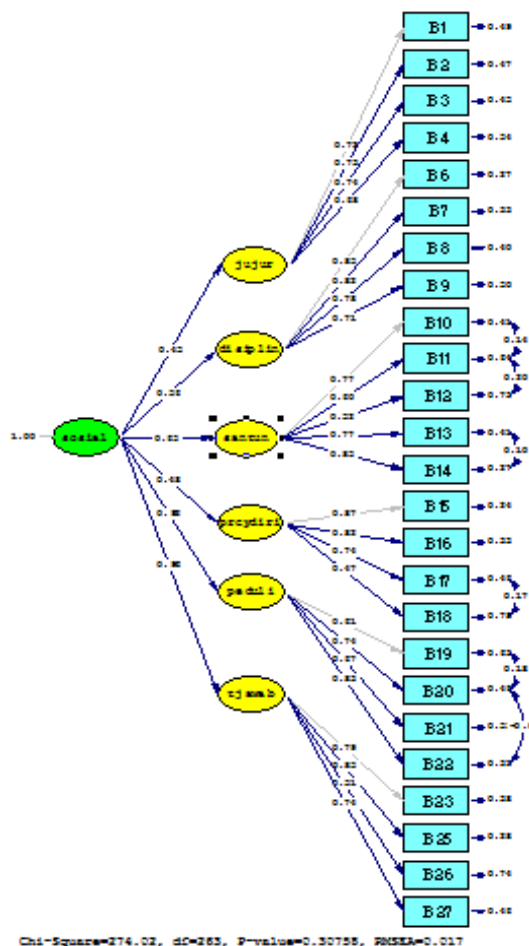


Figure 2. Output Lisrel Path Diagram Factor Value Loading Instrument (without B5 and B24)

The result of data analysis shows that the construct used to form a research model, in the process of confirmatory factor analysis has met the criteria of goodness of fit (Widarjono, 2010) that has been determined. Based on the output on Lisrel it appears that (1) Root Man Squares of Approximation (RMSEA) shows a value of 0.017 (≤ 0.08); (2) Godnees of Fit Index (GFI) shows a value of 0.87, (3) adjusted Goodness of Fit Index (AGFI) shows a value of 0.84 (≥ 0.80); and (4) Root Mean Residual Squares (RMSR)

shows a value of 0.041 (≤ 0.05). Based on these results it can be concluded that the model is considered worth (goodness of fit).

The result of factor loading coefficient value analysis can be seen in Table 2.

Table 2. Estimation of Factor Loading Value Estimation

No	Laten Variable	Indicator	Estimation	Error Standar	Criteria (sig/not)
1	Honest	Item 1	0,72	0,48	sig
		Item 2	0,73	0,47	sig
		Item 3	0,74	0,42	sig
		Item 4	0,68	0,24	sig
2	Disciplin	Item 6	0,85	0,27	sig
		Item 7	0,82	0,22	sig
		Item 8	0,78	0,40	sig
		Item 9	0,71	0,20	sig
3	Courtesy	Item 10	0,6	0,24	sig
		Item 11	0,52	0,27	sig
		Item 12	0,77	0,48	sig
		Item 13	0,85	0,27	sig
4	Self Confident	Item 14	0,87	0,24	sig
		Item 15	0,82	0,22	sig
		Item 16	0,74	0,47	sig
		Item 17	0,47	0,72	sig
5	Caring	Item 18	0,61	0,23	sig
		Item 19	0,74	0,42	sig
		Item 20	0,67	0,21	sig
		Item 21	0,83	0,22	sig
6	Responsibility	Item 22	0,79	0,28	sig
		Item 23	0,85	0,26	sig
		Item 24	0,51	0,28	sig
		Item 25	0,74	0,42	sig
		Item 26			
		Item 27			
		Item 28			
		Item 29			

Item 26
Item 27

Table 2 shows that the influence of latent variables of honest, disciplined, polite, confident, caring, and responsibility towards all 25 indicator points is statistically significant. The Godness of Fit Index test results have met the criteria, showing all latent variables theoretically can be measured by their respective indicators so that the proposed theoretical model is declared in accordance with the empirical data.

Reliability is calculated by construct reliability formula and variance extract. Can be seen in Table 3.

Table 3. Construct Reliability Value

No	Aspect	Indicator	Construct Reliability Value	Criteria
1	Honest	Item 1	0,938	Reliabel
		Item 2		
		Item 3		
		Item 4		
2	Discipline	Item 6	0,869	Reliabel
		Item 7		
		Item 8		
		Item 9		
3	Courtenous	Item 10	0,834	Reliabel
		Item 11		
		Item 12		
		Item 13		
4	Self Confident	Item 14	0,823	Reliabel
		Item 15		
		Item 16		
		Item 17		
5	Caring	Item 18	0,810	Reliabel
		Item 19		
		Item 20		
		Item 21		
6	Responsibility	Item 22	0,818	Reliabel
		Item 23		
		Item 24		
		Item 25		
		Item 26		
		Item 27		

All aspects of social behaviours have a value > 0.70, namely: honest (0.938 > 0.70), discipline (0.869 > 0.70), Courtenous (0.834 > 0.70), self confidence (0.823 > 0,70), cares

(0.810 > 0.70), and responsibility (0.818 > 0.70). This calculation can be understood that theoretically the results of research in accordance with empirical data and have a reliable model.

Therefore, with this result it can be said that the items of social behaviours instrument is a good point and can be used as a measure of students' social behaviours and instrument reliability coefficients are in the high category so that the instrument can be regarded as a consistent or reliable instrument. This is in line with that revealed by the dragon that the reliability coefficient of 0.50 is sufficient enough to be accepted as a good reliability (Khumaedi, 2012).

Likewise with the results of research from Rochmiyati (2013) which shows that the instrument of peer assessment of social skills developed valid and reliable and through the validity of constructs that meet the requirements of the good category. In addition to validity, the reliability coefficient generated on the developed instrument is in the high and very high category. For the reliability coefficient generated on the developed instrument is in high and very high category. This is in line with research Nur, et al (2015) authentic assessment instrument there is a score test results that get that the coefficient of reliability is in the high category (> 0.6) so that the instrument is declared reliable.

Likewise with the research Nurhadi (2015) the validity of the instrument is very high and the results of the reliability of the behaviours assessment instrument into the category of instrument reliability is very high. Likewise Hayatun research (2015) indicates the validity of sufficient instruments and the coefficient of reliability of the behaviours instrument entered in the category of being.

CONCLUSION

Based on the results and discussion of the research that has been described, it can be concluded that the instrument developed has

several characteristics, namely the use of instruments developed more specifically used by teachers to assess students' social behaviours, developed instruments used by students for peer assessment and developed instruments used side by side observation sheet from the teacher. The validity of the contents of the developed instrument is based on expert judgment and analyzed using Aiken's V formula. The results show that all items of statement from every aspect of social behaviours are valid because they have validity value > 0.3. While the construct validity is based on field trial results and analyzed using confirmatory factor analysis. The results of the analysis show that the instruments of social behaviours are 25 points of valid statements of 35 statements, namely 4 honest statements, 4 statements of discipline, 5 statements of polite behaviours, 4 statements of self-confidence, 4 statements of caring behaviours, and 5 statements of behaviours answer. The reliability of the developed instrument is based on the results of field trials and analyzed using onstruct validity. The results show that from all aspects of social behaviours have a high reliability coefficient. So that the developed instrument is valid and reliable.

REFERENCES

- Agus, Widarjono. (2010). Analisis Statistika Multivariat Terapan. Yogyakarta: UPP STIM. YKPN.
- Asmarawati, Endah, Riyadi & Imam Sujadi (2016). Proses Integrasi Sikap Sosial dan Spiritual Dalam Pembelajaran matematika Pada Siswa Kelas VII SMP Negeri Di Kecamatan Purwodadi. Jurnal elektronik Pembelajaran Matematika Vol.4, No.1, Maret. 58-69.
- Azarnoosh, Maryam. (2013). "Peer assessment in an EFL context: attitudes and friendship bias". Language Testing in Asia, 3(11):1-10.
- Dinar, Al Widian, Mahanani, Isnawati. (2016). Pengembangan Instrumen Penilaian

- Autentik Dalam Pembelajaran Biologi materi Fungi kelas X SMA/MA. Berkala Ilmiah Pendidikan Biologi (BioEdu). Vol.5 No.3, September.
- Gulikers, Judith, et.al. (2009). *The Power of Assessment in teacher Education*. Springer Science+Business Media B.V.
- Hayatun Nufus, Sabrina, Abdul Gani & Suhendrayatna. (2017). Pengembangan Instrumen Sikap berbasis Kurikulum 2013 Pada Pembelajaran Kimia SMA. *Jurnal pendidikan Sains Indonesia*, Vo. 5, No.01. 44-51.
- Khumaedi. (2012). Reliabilitas Instrumen Penelitian Pendidikan. *Jurnal Pendidikan Teknik Mesin*, 12(6), 29.
- Khuriyah. (2003). Pengembangan Instrumen Evaluasi ranah Afektif Untuk Pendidikan Agama Islam. *Jurnal Penelitian dan Evaluasi*, Nomor 6 Tahun V.
- Nur Anggraeni, Nadya, Sriyono, Nur Ngazizah. (2015). Pengembangan Instrumen Penilaian Autentik untuk mengukur Sikap Sosial peserta Didik Sma kelas X pada Pembelajaran Fisika. *Radiasi*. Vol.7, No2, September.
- Nurhadi. (2015). Pengembangan Instrumen Penilaian Sikap Spiritual dan Sosial Pada Pembelajaran IPA Terpadu. *Jurnal Pembelajaran Fisika Universitas Negeri Lampung*.
- Rochmiyati. (2013). Model Peer Assessment Pada Pembelajaran Kolaboratif Elaborasi IPS Terpadu Di Sekolah Menengah Pertama. *Jurnal Penelitian dan Evaluasi Pendidikan Tahun 17*, Nomor 13.
- Rusilowati. (2013). Pengembangan Instrumen Non Tes. *Makalah*. Seminar Nasional Evaluasi Pendidikan I di Universitas Negeri Semarang.
- Setiadi, Hari. (2016). Pelaksanaan Penilaian Pada Kurikulum 2013. *Jurnal Penelitian dan Evaluasi Pendidikan Volume 20*, No2, Desember.
- Sugiyono. (2010). *Metode Penelitian Pendidikan Pendekatan Kuantitatif, Kualitatif, dan R&D*. Bandung: Alfabeta.
- Widoyoko, E.P. (2014). *Evaluasi Program Pembelajaran*. Yogyakarta: Pustaka Pelajar.
- ZevenBergen, Robyn. (2001). Peer Assessment of Student Constructed Posters: Assessment Alternatives In Pre Service Mathematics Education. *Journal of Mathematics Teacher Education* 4: 95–113.