Development and Validation of Student's Responsibility Scale on Mathematics Learning Using Subject Scaling Model

Kana Hidayati
Graduate School of Yogyakarta State University, Indonesia, *kana@uny.ac.id*

Budiyono
Sebelas Maret University, Indonesia, *budiyono@staff.uns.ac.id*

Sugiman
Yogyakarta State University, Indonesia*, sugiman@uny.ac.id*

The character of responsibility is an important character that Indonesian students must possess. The attempts to strengthen the character of student responsibilities require in-depth information about the actual state of student's responsibility. The use of relevant, valid, and reliable instruments is necessary to obtain accurate and accountable information about the conditions of student's responsibility. A scale for measuring student's responsibility in mathematics learning was developed through this study. The scale was developed by using subject scaling models and tested on junior high school students in Yogyakarta, Indonesia. The first trial involved 688 students and a second trial involved 696 students. Based on this study, the result of Student's Responsibility Scale on Mathematics Learning (SRSML) consisted of 14 items in the form of subject scaling which theoretically and empirically proved to be of good quality. The scale has good content validity supported by V Aiken index of 0.88-0.92. The scale has good construct validity. Scale reliability is also categorized as good with Omega coefficient of 0.89. SRSML has a lot of potential to be used both in the character assessment activities of students in learning mathematics to strengthen character education and educational research, especially related to the student's responsibility character.

Keywords: student's responsibility, scale development, subject scaling, validity, reliability.

INTRODUCTION

Indonesian National Education System strongly emphasizes education that is able to make Indonesian people not only have intellectual intelligence but also noble character with good spirituality, personality, and self-control. Therefore, character education is

highly emphasized in Indonesian education system and continues to undergo reforms following the development of the era.

The implementation of Curriculum 2013 in Indonesia that emphasizes character education and the establishment of the Character Education Strengthening program in schools are several forms of government commitment in order to strengthen the noble character that Indonesian students must possess. Minister of Education and Culture mentioned that the portion of science education for junior high school education is 40% and character education is 60% (Maharani, Republika: September 6, 2016).

Literature Review/Theoretical Background

Good character education contains knowledge which further raises the commitment to be a good person and finally really embodied in doing good behavior. This is in line with Lickona (1992) who stated that good character education emphasizes on three groups of good character namely moral knowing, moral feeling, and moral action. According to the taxonomy of Krathwohl et al. (1964), this situation corresponds to the affective domain. It becomes the benchmark of attitude competence in the Curriculum 2013 namely receiving, responding, valuing, organization, and characterization by value.

Assessment of affective aspect in Curriculum 2013 in Indonesia really emphasizes on character education. Therefore, the assessment of attitude aspects in mathematics learning should also include measurement and assessment of the good characters that Indonesian students must have. If the assessment of attitude aspects in mathematics learning also measure the good characters of the students, the results will be very helpful for teachers to take steps in an effort to strengthen their good characters. Kumaidi (2014) suggests that the Curriculum 2013 requires teachers not only assess the mastery of teaching materials but also include the development of affection such as faith, honesty, discipline, and other good characters.

One of the emphasized good characters that Indonesian students must have is responsibility. Strengthening the responsibility character of the students in learning mathematics certainly requires in-depth information about the actual conditions of the responsibility character that the students have. Therefore, measuring the character of responsibility in mathematics learning is a must. It should be done by using appropriate measurement instrument to help teachers in assessing students' attitude or affective aspects. Popham (2009) suggests that students' affective assessments should be done during school learning activities as it can contribute to student behavior after graduation. This means that the measurement and assessment of the responsibility character of students during the school greatly affect the behavior of student responsibility after graduation.

The notion of responsibility according to Lickona (1992) is to carry out obligations or work both within the family, school, or workplace by giving the best and done wholeheartedly. This is in line with the definition of responsibility according to Miller (2005) who states that responsibility is readiness to bear the consequences of behavior or actions taken and do the right thing at the right time so that others trust. This shows

that a responsible person will be firm in making decisions and be ready to bear the consequences or risks that exist. A responsible person will always be ready and dare to bear everything or in other words dare to answer and bear the consequences.

The responsibilities according to Zuchdi (2011) are outlined as follows: (a) fulfilling the duties of self, (b) trustworthy, (c) self-controlling, (d) persistent, (e) preparing for the best, (f) on time, (g) self-discipline, and (h) can work with friends in a team. According to Kemendikbud (2015), responsibility is defined as the behavior of a person to carry out the duties and obligations that he should do to self, environment (nature, social and culture), society, country and Almighty God. Behaviors that reflect the responsibility character may be: (a) carrying out the individual's duties well, (b) taking risks for any action taken, (c) not accusing others without accurate evidence, (d) returning borrowed items, (e) acknowledge and apologize for the wrong doing, (f) keep the promise, (g) not blame others for their own misdeeds, and (h) willingness to carry out what was once said.

Based on the various opinions above, it can be concluded that the responsibility is the attitude of a person who is firm in making decisions and ready and dare to bear the risks or consequences for each decision. Based on this definition, it can be operationally defined that responsibility is the behavior to carry out the duties and obligations as they should be in good and wholeheartedness (commitment to task), ready to assume any risks or consequences (gentlemanliness), and acknowledge errors and able to complete the task until completed to completion (discipline).

The measurement and assessment of attitudes related to the character of students' responsibilities in learning mathematics are generally done by observation, interview, self-assessment, and assessment among friends. The instrument forms used are in the form of observation sheet, interview guide, and questionnaire in the form of scaling response. Character measurement is basically a measurement of individual personality so it requires instruments that can really distinguish personality between individuals. Therefore, the use of scale-shaped scaling subjects that is more specific and detailed in measuring the character of students will be more useful. However, until now there are not many studies on the scale development for measuring the character of the student's responsibility, especially in mathematics learning by using subject scaling model.

Several studies on the scale of responsibilities that have been carried out include: Singg & Ader (2001) who developed Student Personal Responsibility Scale-10 (SPRS-10) to measure student responsibilities in daily life; Akbay, Capri, & Gunduz (2013) who developed an academic responsibility scale using Likert scale with four choices of answers consisting of 25 items; Amanda et al. (2016) who developed personal responsibility scale for adolescents containing three factors and arranged in 15 items; and Eristi (2017) who developed a learning responsibility scale with four learning behavior-forming factors and consisted of 28 items. The various scales that have been developed are using response scaling models.

Scale development to measure and assess the character of student responsibilities using a scaling model of the subject which is directly related to the material in mathematics is

still very rare. In fact, scale-shaped scaling of the subject has a high practical value and widely used the designers of psychological scale (Azwar, 2005). It is also as suggested by Punyanunt (2017) that a subject-centered scale which is also called as the scale of individual differences is a perfect tool for measuring one's personality.

Research Question

This study aims to develop a scale to measure student's responsibility on mathematics learning using subject scaling model. This scale is tested for its quality in a theoretically and empirically. Theoretical quality-scale evidence is obtained through content validity with expert judgment and the empirical evidence of scale quality is viewed from construct validity and reliability (Bolanriwa, 2015).

The research questions of the study were listed as follows.

(1) How is the validity of the content of the scale to measure the character of student responsibility in mathematics learning?
(2) How is the validity of scale constructs to measure the character of student responsibility in mathematics learning?
(3) How is the scale reliability to measure the character of student responsibility in mathematics learning?

METHOD

This research is a development research conducted with several steps namely: (1) preparing initial scale, (2) trial, and (3) interpretation of trial result. Arranging the initial scale is the initial activity of scaling up which is further discussed in the Focus Group Discussion (FGD) and followed up with validation through expert judgment. The final result of validation becomes evidence of the validity of the contents of the developed scale. The next stage is a trial and follow up interpretation of test results that determine the evidence of quality from the scale related to the validity and reliability of the construct. Then it is determined whether or not the revision is needed and finally tests it again.

The scale for measuring the character of student responsibility in mathematics learning in this study is arranged in the form of subject scaling with three answer choices. The items are presented in the form of a narrative sentence about conditions encountered by students in learning mathematics. The choice of answers are also presented in the form of a sentence that states the various responses of students related to the conditions encountered as stated in the question items. The scoring of the answer options are: (a) score 2 for the answer option that most indicates the character of the responsibility, (b) score 1 for the answer option that indicates less character responsibility, and (c) score 0 for the answer option that does not indicate the character of responsibility.

FGD activities involve 4 mathematics teachers of Junior High School and 3 lecturers of Mathematics education. Scale validation with expert judgment involves six experts reviewing the scale in terms of its content: (a) the suitability of the item with the indicator, (b) the language used, (c) the unbiased statement of the item, (d) the clarity of the answer option, and (e) the precision of scoring guidelines. Experts also assess the

appropriateness of items with indicators in the form of Likert scale with five answer choices. In addition to qualitative assessment of the experts, the evidence of content validity is also supported by the content validity index using V Aiken. Determination of validity of an item is done by comparing the value of V calculation results with value in the V_{table}. The V_{table} value is the minimum value of the content validity index based on the number of experts in V_{table} in Aiken (1985). The number of raters in this study are six and the number of ratings are five, so the minimum content validity index based on table V Aiken is 0.79.

Participants

The trial was conducted on junior high school students of class VII in province of Yogyakarta that have implemented Curriculum 2013. School selection is done by purposive cluster sampling technique, it is conducted in private and state junior high schools which represent high quality, medium, and developing school. Determination of the quality of schools in this study is based on the results of the National Examination. The schools selected as pilot sites are SMP N 5 Yogyakarta, SMP IT Abu Bakar Yogyakarta, and SMP Muhammadiyah 2 Yogyakarta.

Data Collection Procedures

The data of the trial results in this study are quantitative data. The data are obtained based on the students' scores on the developed scale to measure the character of the students' responsibility in mathematics lesson. The technique of data collection is done by directly conveying the scale to the students.

Data Analysis

To obtain evidence of construct validity, factor analysis using Exploratory Factor Analysis (EFA) and Confirmatory Factor Analysis (CFA) were used. This technique is also used by Bashir and Bala (2018). This study uses the EFA approach to investigate the factors that shape the responsibility constructs and the CFA approach to confirm these factors in shaping the construct of responsibility.

Using the EFA approach, the Kaiser-Meyer-Olkin Measure of Sampling Adequacy (KMO MSA) and Barlett Sphericity tests are used to determine whether the items' scale matched the factor analysis or not. The match criteria are a minimum MSA KMO score of 0.50 and statistically significant Barlett Sphericity test results (Hair et al., 2010). This study extracts factors using Principal Component Analysis (PCA) and rotation methods to reduce factors using the varimax method. EFA in this study is conducted using SPSS version 20.0.

In CFA approach, data analysis begins with the requirements analysis test to determine whether the data already meet the requirements to be analyzed by CFA technique, which is required to test the model by using the joint multivariate normal distribution. The criteria used is if p-value>0.05 then the distribution is normal, and if p-value≤0.05 then the distribution is not normal (Yamin & Kurniawan, 2009). After the requirements analysis test, data analysis is performed to verify the validity of scale constructs through

second order CFA with the help of Lisrel program version 8.51. The criteria for a valid indicator in representing the construct are if t-value>1.96 and the value of Standarized Loading Factor (SLF) at least 0.3 (Igbaria et al., 1997; Hair et al., 2010). Referring to these criteria, if any indicator items is not valid then the item must be eliminated and re-analyzed. Based on the re-analysis results after the invalid indicator item is omitted, the next step to be considered is its compatibility with the model. If the model is not fit, then the model will be adjusted by modifying the index as suggested by Lisrel.

Related to the model fit, Suranto, Muhyadi, & Mardapi (2014) suggested that the developed model is considered suitable with field data if it already fulfills two criteria from three criteria which become the parameters of absolute fit measures: Root Mean Square Error of Approximation (RMSEA)≤0.08, Chi square obtained from testing has a probability greater than 0.05 (p>0.05), and Goodness of Fit Index (GFI)≥0.90. Meanwhile, Purnomo (2017) uses good model criteria including RMSEA≤0.05, GFI≥0.95, Adjusted Goodness-of-Fit Index (AGFI) ≥0.95, Normed Fit Index (NFI) ≥0.95, Tucker-Lewis Index (TLI) ≥0.95, and the Comparative Fit Index (CFI) ≥0.95. Specifically, Wijanto (2008) suggests criteria for: a) RMSEA: ≤0.08 (good fit) and ≤0.05(close fit), b) GFI: ≥0.90 (good fit) and 0.80≤GFI<0.90 (marginal fit), c) AGFI: ≥0.90 (good fit) and 0.80≤GFI<0.90 (marginal fit), d) NFI: ≥0.90 (good fit) and 0.80≤GFI<0.90 (marginal fit), e) TLI: ≥0.90 (good fit) and 0.80≤TLI<0.90 (marginal fit), and f) CFI: ≥0.90 (good fit) and 0.80≤GFI<0.90 (marginal fit).

The model fit in this study refers to Garson (2009) which suggests that suitability of model developed through empirical data is at least seen from the fulfillment of two of the three different model fit category sizes namely absolute fit measures, incremental fit measures, and parsimonious fit measures. It also refers to the opinion of Hooper, Couglan, & Mullen (2008) which suggests that the use of Chi square is sensitive to the size of the sample used. Furthermore, Wijanto (2008) mentions the consensus of the researchers who state that Chi square is not the only measure of Goodness of Fit (GOF) and there is no GOF size which exclusively represents the overall model fit.

The reliability estimate is determined by the formula of the McDonald's composite coefficient of reliability or the Omega coefficient (ω). This refers to Widhiarso (2009) who suggest that the reliability that begins with CFA using SEM on multidimensional measurement model can be done with Omega (ω) reliability coefficient. In addition, referring to SEM analysis results, it can also determine reliability estimation for each factor using the coefficient of Construct Reliability/CR (Wijanto, 2008). In this case, reliability coefficient criteria refers to Hair et al. (2010) who stated that reliability coefficient >0.7 is good, while the reliability coefficient of 0.6-0.7 can be accepted as long as the indicator of construct validity of the model is good.

FINDINGS

Based on the results of the initial scale arrangement which is then discussed in the FGD and followed up with validation by the experts, Student's Responsibility Scale on

Mathematics Learning (SRSML) is developed. It consists of 14 items in the form of subject scaling. Indicators of SRSML are presented in Table 1.

Table 1
Indicators of SRSML

Factor	Indicator	Item
A. Commitment to the task	A1. Do every task given by the teacher seriously	T1
	A2. Always ready to do the task from the teacher	T2
	A3. Play an active role to solve problems in every task given by the teacher	T3
	A4. Do the task without being forced.	T4
	A5. Display initiatives in task completion.	T5
B. Gentlemanliness	B1. Willing to admit their own mistakes.	T6
	B2. Refuse to accuse others without accurate evidence.	T7
	B3. Willing to apologize for the mistakes made.	T8
	B4. Willing to resign because of failure toperform the task, if it is the best solution for the public interest	T9
	B5. Willing to be subject to legal sanctions if proven to violate regulation	T10
C. Discipline	C1. Obey the rules in the classroom.	T11
	C2. Obey the rules in group.	T12
	C3. Keeping their promise.	T13
	C4. Returning any items borrowed from others on the promised time	T14

Based on the validation results, the initial SRSML consisting of 14 items in the form of subject scaling has good content validity so it is feasible to use. The feasibility of the developed scale is also supported by the magnitude of the V Aiken index for the suitability between the items with the indicator compared with the value of V_{table} which is 0.79. Based on the calculation results, values of V on the overall items are 0.88-0.92. This indicates that all SRSML items have good content validity in terms of their suitability with the indicators. Therefore, in terms of content, the SRSML developed by this study can be said to measure what should be measured.

The scale that has been declared as eligible by the experts is then tested on the first trial involves 688 students. Based on trial data I, EFA is done to identify the factors that make up the responsibility constructs. Based on SPSS version 20.0, the value of Barlett Test of Sphericity is 965,300 with 0,000 significance level. This shows a significant correlation between variables. The calculation result of KMO MSA is 0.781 which indicates that the adequacy of the sample is good. Those results indicate that factor analysis is feasible. Factor extraction was done using PCA method and continued by factor rotation using varimax method. Based on the extraction results, three factors forming responsibility constructs are obtained. The first factor explains the variance of 20.776%, the second factor is 9.740%, and the third factor is 7.930%. From the results of rotationing and looking at the loading factor of each variable, three groups are obtained: the first factor is named "commitment to the task" and contains items T1, T2, T3, T4, T5; the second factor is named "gentlemanliness" contains T6, T7, T8, T9, T10; and the third factor is named "discipline" contains points T11, T12, T13, T14. Thus, it can be identified that SRSML consists of 3 factors namely commitment to the task, gentlemanliness, and discipline.

For the next step, CFA is conducted to confirm that all three factors really explain the construct of responsibility. The result of preliminary analysis shows that the data is not multivariate normal distribution, so CFA second order analysis is done with Weighted Least Square (WLS) estimation model. Based on Lisrel analysis result from SRSML data on trial I, the value of RMSEA=0.046, Chi square=181.91 with p-value=0.00, GFI=0.98, AGFI=0.98, TLI=0.83, IFI=0.86, and CFI=0.86. Although Chi square criteria is not fulfilled because p-value is less than 0.05, yet other criteria have been met. Therefore, by referring to Hooper, Couglan, & Mullen (2008) and Hair et al. (2010), the suitability of the model can be considered to have been adequately met. The CFA second order results for t-value and Standardized Loading Factor (SLF) on SLSMR are presented as follows.

Table 2
Results of Second Order CFA SRSML on Trial I

Factor	Indicator	Item	Second order CFA	
			t-value	SLF
A	A1	T1	**	0.58
	A2	T2	11.93	0.54
	A3	T3	12.98	0.69
	A4	T4	13.75	0.72
	A5	T5	14.10	0.71
B	B1	T6	**	0.50
	B2	T7	7.56	0.54
	B3	T8	8.30	0.60
	B4	T9	2.28	0.16
	B5	T10	10.14	0.66
C	C1	T11	**	0.42
	C2	T12	6.83	0.44
	C3	T13	6.22	0.38
	C4	T14	6.75	0.61

Table 2 shows that in terms of t-value, all of the items are significant in order to support the construct of responsibility, but there is one item that has a SLF value of less than 0.3 which is item T9. Therefore, since this item represents one indicator it is not eliminated but undergo revision process. After the revision, a second trial involving 696 junior high school students is conducted.

Trial II was conducted on 14 items that have undergone improvement based on trial result I. Before CFA conducted on trial II data, EFA also used to extract factors that build character of student responsibility in mathematics learning. Based on the results of analysis, the Barlett Test of Sphericity value of 1402.209 is obtained with a significance level of 0.000. It means that there is a significant correlation between variables. The calculation of KMO MSA of 0828 indicates that the adequacy of the sample is good. Factor extraction was done using PCA method and 3 factors were produced. The first factor explains variance of 24.647%, second factor 10.789%, and third factor 7.207%. Rotation factor to maximize the grouping of variables is done by varimax method. Based on the result of rotationing and paying attention to factor loading of each

variable, there are three constructing factors of students responsibility in mathematics learning namely: the first factor (commitment to the task) which contains items T1, T2, T3, T4, T5, the second factor (gentlemanliness) which contains items T6, T7, T8, T9, T10, and the third factor (discipline) which contains items T11, T12, T13, T14.

After that, CFA is conducted to confirm that all three factors really explain the SRSML construct. The result of data analysis in trial II shows that the analyzed data is not multivariate normal distribution so that CFA second order analysis is done with Weighted Least Square (WLS) estimation model. Based on the preliminary analysis on trial II data, no item has SLF value less than 0.3. It means that no items are eliminated. Further analysis is performed on all items with respect to Lisrel's suggested index modification to improve model match. Here is the SRSML measurement model with modified model as suggested by Lisrel.

The index modification in this study refers to Wijanto (2008) which suggests two suggestions in index modification: (1) adding path between latent variables with observed variables, and (2) adding covariance between two errors. The addition of the track is done if it has a strong enough theoretical support, but if not it should not be done. This study uses the second suggestion of adding covariance between two errors in the measurement model of the same latent variable.
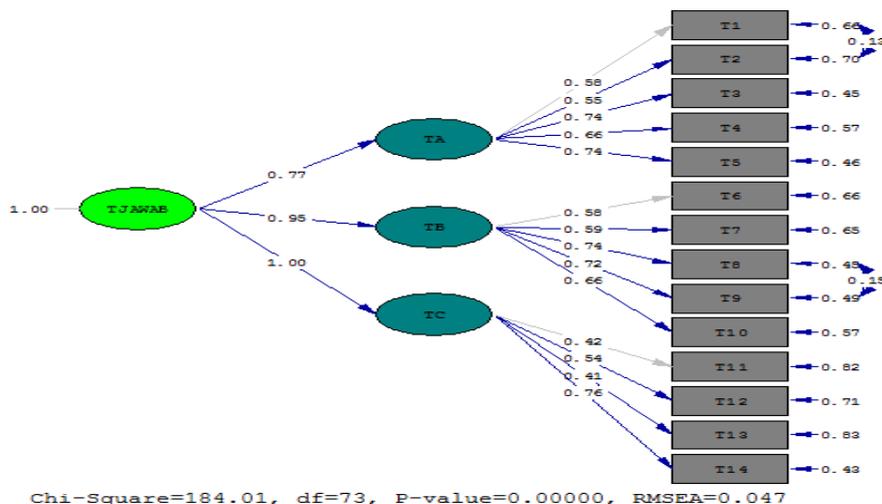

Figure 1
Confirmatory Factor Analysis Model of SRSML

As presented on Figure 1, the modification done by adding covariance between two errors on T1 and T2, T8 and T9. The researcher's consideration is that each of them represents the same latent variable and decreases the Chi square value. Additionally, the addition of covariance between two errors on T1 and T2 and T8 and T9 is reasonable or can be supported by a strong reason. The content of indicator T1 is to implement

seriously every task given by the teacher and T2 is always conditioned to be ready to do the task from the teacher. The errors that occur in T1 are obviously very influential on the T2 indicator. The content of T8 indicator is willing to apologize for the mistakes made and T9 is willing to resign for failing to perform the task, if that is the best solution for the public interest. Error that occurred on the T8 is also clearly very influential on the T9 indicator. This shows that it is reasonable to add covariance between two errors on T1 and T2 and T8 and T9.

Based on final analysis result after modification, the value obtained are RMSEA=0.047, Chi square=184.01 with p-value=0.00, GFI=0.98, and AGFI=0.98, TLI=0.89, IFI= 0.91, and CFI=0.91. As the results of the analysis on trial I, in this case although Chi square criteria is not fulfilled because p-value is less than 0.05, but other criteria have been met. Therefore, according to Hooper, Couglan, & Mullen (2008) and Hair et al. (2009), the suitability of the model can be considered to have been adequately met. The CFA second order results for t-value, Standardized Loading Factor (SLF), and the calculation results of Construct Reliability (CR) on SLSMR for trial II are presented as follows.

Table 3
Results of Second Order CFA SRSML on Trial II

Factor	Indicator	Item	Second order CFA		CR
			SLF	t-value	
A	A1	T1	0.58	**	0,79
	A2	T2	0.55	12.13	
	A3	T3	0.74	13.27	
	A4	T4	0.66	12.30	
	A5	T5	0.74	12.60	
B	B1	T6	0.58	**	0,79
	B2	T7	0.59	9.77	
	B3	T8	0.74	12.25	
	B4	T9	0.72	12.49	
	B5	T10	0.66	11.93	
C	C1	T11	0.42	**	0,91
	C2	T12	0.54	7.99	
	C3	T13	0.41	6.98	
	C4	T14	0.76	8.35	

Table 3 shows that in terms of t-value, all SRSML items are significant in order to support constructs of responsibility. Judging from the SLF value, no item has a SLF value less than 0.3. This means that all items in the SRSML meet the criteria of good construct validity. Based on calculation results, it is obtained that the CR coefficient for factor A is 0.79, factor B is 0.79, factor C is 0.91, and Omega reliability coefficient is 0.89. This shows that SRSML has good reliability. That is, SRSML is said to be reliable or consistent.

DISCUSSION AND CONCLUSION

This study produces a Student's Responsibility Scale on Mathematics Learning (SRSML) which is composed of three constructing factors of responsibility, namely:

commitment to task, sportsmanship, and discipline. The resulting SRSML consists of 14 items in the form of subject scaling that theoretically and empirically proved to be of good quality. Theoretically, it is proved by the validity of the content scale that indicates that the content is good according to the experts judgement and the whole item has good conformity with the indicator. This is supported by the magnitude of V Aiken index of 0.88-0.92. Empirically, from the validity of the construct, all items of scale have good support in the establishment of responsibilities construct. Reliability is also belongs to good category with the magnitude of CR coefficient for factor A is 0.79, factor B is 0.79, factor C is 0.91, and Omega reliability coefficient is 0.89.

SRSML items in the form of subject scaling have the advantage of being able to reveal the character of student responsibility more specifically and deeply. The items and choices of answers in the SRSML are presented in a narrative form so as to better reflect the actual student condition. This is consistent with Punyanunt (2017) that a subject-centered scale or the scale of individual differences is well suited for measuring one's personality. In addition, the procedure has a high practical value and is widely used by designers of psychological scales (Azwar, 2005).

Some examples of SRSML items in the form of subject scaling generated in this study are T3, T5, T8, and T14. The items are the items that have the highest SLF value in each factor. The following indicators are represented by each item.
Item T3: Play an active role to solve problems in every task given by the teacher (SLF= 0.74).
Item T5: Display initiatives in task completion (SLF=0.74).
Item T8: Willing to apologize for the mistakes made (SLF=0.74).
Item T14: Returning any items borrowed from others on the promised time (SLF=0.76).

The SRSML items in the form of subject scaling in this study were compiled with three options of answer. The scoring guidelines used are: score 2 for answer choices that most indicate the character of responsibility, score 1 for answer choices that indicate less character responsibility, and score 0 for answer options that do not indicate the character of responsibility. The following descriptions are T3, T5, T8, and T14 items and their scoring guidelines.
Item T3:

> During math classes, group discussions often take place to do the activity sheets provided by the teacher, then what you do is ...
> a. *Trying to always play an active role in the discussion to solve the problems in each activity sheet provided by the teacher.*
> b. *If there is a student who excels in Math in the group, I just follow his opinion.*
> c. *I prefer to listen to the opinion of other friends before trying to participate in the discussion in the group.*

The scoring guide for item T3 is score 2 for option a, score 0 for option b, and score 1 for option c.
Item T5:

> One day, the math teacher was unable to attend and gave the group tasks to collect data on the types of vehicles in the school parking lot, so what you do is ...

    a. Take an active role in submitting ideas in the completion of the task.
    b. Waiting the direction from group leader.
    c. Works only on the assigned parts.

The scoring guide for item T5 is score 2 for option a, score 0 for option b, and score 1 for option c.

Item T8:

    One day, your friend asked for your help explaining how to determine the center angle of making a pie chart. Inadvertently you are wrong to explain, but your friends do not know and follow the way you explain. After seeing the results of his work, you just realized there is a wrong step, then you:
    a. Apologize if the friend asked you about it.
    b. Just be quiet because it is not intentional and the friend also did not know it.
    c. Instantly apologize, explain the wrong move and show the correct step.

The scoring guide for item T8 is score 1 for option a, score 0 for option b, and score 2 for option c.

Item T14:

    When the mathematics lesson discusses the pie chart, you forget to bring the compass. You see your friend carrying two very good compass and looks like new. You are allowed to borrow one of those compass. When the mathematics lesson is finished, then you:
    a. Keep borrowing the compass because you feel you still need it and will return it if asked.
    b. Not immediately return the compass and expect him to give you the compass because he has two compasses.
    c. Immediately return the compass and say thanks to your friends' help.

The scoring guide for item T14 is score 1 for option a, score 0 for option b, and score 2 for option c.

Through the form of these items, students are expected to provide answers that really indicate the actual state of character of student responsibilities when faced with certain situations as stated in the items. Referring to Lickona (1992), the SRSML that has been produced in this study can be said to measure the character at the morality level of feeling. That is, the revealed character conditions reflect the commitment of students to behave with responsibility. This is also consistent with the affective domain taxonomy of Krathwohl which is up to the organization stage. This means that students have established a consistent value of system within themselves.

The SRSML produced by this study is highly relevant to character education in Indonesia. This scale is very useful in adding references for teachers, especially math teachers in doing affective assessment of students. In addition, this scale can also be used by teachers and education practitioners as a reference for conducting research such as classroom action research or experiments. For other researchers, this scale can be a reference related to the form of psychological scale in subject scaling. This is because such scales are still very rarely developed especially for affective assessment in learning Mathematics.

Calculating the score of the scale in this research is still done manually, so it takes a relatively long time to get the data. For further development, this instruments can be made in the form of software or presented online so scoring can be done more easily and quickly. This is in line with Leachy (2012) who argues that technological developments should be able to be utilized by teachers in assessment activities because technology can be designed and utilized in order to provide a quick response to student work. Further trial activities also need to be conducted in other areas with larger and heterogeneous respondents, thus further expanding the generalization and increasing evidence of the scale quality produced by this study. Further development of this scale can also be carried out on the development of scale to measure other main character which is the priority of character building activities in the education system in Indonesia.

REFERENCES

Aiken, L. R. (1985). Three coefficients for analyzing the reliability and validity of ratings. *Educational and Psychological Measurement*, 45, 131-142.

Akbay, S. E., Capri, B., Gunduz, B. (2013). Development of the academic responsibility scale (ARS): a validity and reliability study. *International Journal of Academic Research* Part B; 2013; 5(4), 440-446. DOI: 10.7813/2075-4124.2013/5-4/B.64.

Amanda, Mergier, Paul, & Shield. (2016). Development of the personal responsibility scale for adolescents. *Journal of Adolescence,* 51, 50-57. https://doi.org/10.1016/j.adolescence.2016.05.011

Azwar, S. (2005). *Penyusunan skala psikologi*. Edisi 1. Yogyakarta: Pustaka Pelajar.
Bashir, H., & Bala, R. (2018). Development and Validation of Academic Dishonesty Scale (ADS): Presenting a Multidimensional Scale. *International Journal of Instruction*, 11(2), 57-74. https://doi.org/10.12973/iji.2018.1125a

Bolarinwa O. A. (2015). Principles and methods of validity and reliability testing of questionnaires used in social and health science researches. *Nigerian Postgraduate Medical Journal,* 22, 195-201. http://www.npmj.org/text.asp?2015/22/4/195/173959

Eristi, B. (2017). Development of a learning responsibility scale. *Education Electronic Journal of Science and Mathematics Education,* 11(1), 481-503.

Garson, G. D. (2009). *Overview structural equation modeling*, http://faculty.chass.ncsu.edu/garson/PA765/structur.htm

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). *Multivariate data analysis* (7th ed). Prentice Hall [versi elektronik].

Hooper, D., Couglan, J., Mullen, M. R. (2008). Structural equation modelling: Guidelines for determining model fit. *Electronic Journal of Business Research Methods,* 6(1), 53-60.

Igbaria, M., Zinatelli, N., Cragg, P., & Cavaye, A. L. M. (1997). Personal computing acceptable factors in small firms: A Structural Equation Model. *MIS Quarterly*, September, 279-299.

Kemendikbud. (2015). *Panduan Penilaian untuk Sekolah Menengah Pertama*. Jakarta: Direktorat Pembinaan Sekolah Menengah Pertama.

Krathwohl, D. R., Bloom, B. S., & Masia, B. B. (1964). *Taxonomy of educational objectives: Handbook II: Affective Domain*. New York: David McKay.

Kumaidi. (Maret 2014). Implementasi penilaian autentik dalam pembelajaran di kelas. *Makalah* disajikan dalam Seminar Nasional Implementasi Penilaian dan Pelaksanaan Kurikulum 2013, di Universitas Negeri Jakarta.

Leachy, G. (2012). 'QR code in mathematics classrooms'. *Mathematics Teaching Issue,* 235: 27-29. Derby UK: The Association of Teacher of Mathematics.

Lickona, T. (1992). *Educating for character*. New York: Bantam Books.

Maharani, E. (12 Juni 2016). DIY targetkan 25 persen sekolah terapkan K-13. *Republika.co.id.*

Miller, J. (2005). *10-minute life lessons for kids.* Harper Collins Publishers.

Popham, W. J. (2009). *Test better, teach better: the instructional role of assessment.* United States of America: ASCD (Association for Supervision and Curriculum Development)

Punyanunt, N. M., (2017). Response options for scales: does it mater what words you use?. *Journal of Media Critiques*, 3(9). doi: 10.17349/jmc117103, http://www.mediacritiques.net

Purnomo, Y. W. (2017). A scale for measuring teachers mathematics-related beliefes: A validity and reliability study. *International Journal of Instruction*. 10(3). 143-158. https://doi.org/10.12973/iji.2017.10310a.

Singg, S., & Ader, J. (2001). Development of the student personal responsibility scale-10. *Social Behavior and Personality: An international journal*, *29*, 331-336. doi: https://doi.org/10.2224/sbp.2001.29.4.331

Suranto, Muhyadi, & Mardapi, D. (2014). Pengembangan instrumen evaluasi uji kompetensi keahlian (UKK) administrasi perkantoran di SMK. *Jurnal Penelitian dan Evaluasi Pendidikan,* 18(1), 98-114.

Widhiarso, W. (2009). Koefisien reliabilitas pada pengukuran kepribadian yang bersifat multidimensi. *Psikobuana,* 1(1), 39-48.

Wijanto, S.H. (2008). *Structural equation modeling dengan lisrel 8.8.* Yogyakarta: Graha Ilmu.

Yamin, S. & Kurniawan, H. (2009). *Structural equation modeling: Belajar lebih mudah teknik analisis data kuesioner dengan Lisrel-PLS.* Jakarta: Salemba Infotek.

Zuchdi, D. (2011). *Pendidikan karakter dalam perspektif teori dan praktik.* Yogyakarta: UNY Press.