

The Real Development of Authentic Assessment Based on Characters of Primary Students

¹Deny Setiawan and ²Harun Sitompul

¹Department of Education, State University of Medan (Unimed), Indonesia

²Department of Educational Technology, State University of Medan (Unimed), Medan, Indonesia

Abstract: The development of this research is intended to see a character-based on authentic assessment instruments at characters and skills of students in primary school. Subjects of this study are primary school teacher primary school students seven experts for validation of assessment instruments with the following criteria mathematician teacher in primary school science teacher in primary school social science teacher in primary school Indonesian language teacher in primary school civics teacher in primary school assessors expert and psychologists. The method used in this study is development research on the type of formative research. Its development is focused on two phases: preliminary phase and formative evaluation phase to include self evaluation, prototyping (expert reviews, one-to-one and small group) and field test. Data are collected through questionnaires and field trials. Expert validation results show that character-based on assessment instruments are valid, good and decent to see student's character with a note it should be revised for the sub-components which having not good categorized. The results of field trials (small scale and large scale) indicates that the character-based on assessment instruments developed attitudes have good validity and have very high reliability. Thus, based on the results of expert validation and field trials, it can be concluded that the attitude of character-based on assessment instruments are valid and have very high reliability category, so the instrument is eligible to use in measuring the achievement of the student character value.

Key words: Authentic assesment, character, primary school, students, Indonesia

INTRODUCTION

Recently, character building at primary school is very necessary. Character as a personality of someone is formed from the internalization of virtues which are believed and used as a basis of way of thinking behave and act. Here virtue consists of a number of moral values and norms such as honesty, courage to act, trustworthy, respect for others, self-discipline, independent, hard working and creative. National identity seems deprived from its roots, so the character becomes a hot topic to be discussed again lately. At the article number I of National Education Act, 2003 states that among the objectives of national education are to develop potential learners to have the intelligence, personality and noble character. Mandate of Education Law, 2003 is not only intended to establish education in Indonesia by producing intelligent students but also having personality or character, so that it will be born generation who grows up with a character consisting of noble values of the nation and religion.

The importance in character building for primary school students should be supported by possessing a

complete learning media which can support the strengthening of student's character. One of learning media needed by students is an authentic assessment oriented in character. To reach the goal of character needed, done indirectly through various learning activities undertaken, teachers are expected to make an assessment for students directly to see a certain characters value reached by student (Majid, 2014).

But the fact, most teachers do not understand much about the concept and implementation of authentic assessment focused on character oriented. It is informed by Rusilowati in her survey against 20 from 23 teachers of SMP 21 Semarang. The result is 87% teachers are having some difficulties in doing authentic assessment. From 23 respondents as many as 87% teachers have difficulty in understanding how to vote and 70% teachers have difficulty in making observation instrument.

The above data is also supported by preliminary observations on the elementary school in Medan. Almost all of the teachers have difficulties in doing the assessment. It happens because many teachers do not understand how to do authentic assessment, especially in

assessing the achievement of student's character. There are 5 elementary schools which have been observed in Medan and all of them expresses some difficulties in implementing authentic assessment to measure the achievement of a certain character values on students. (Hadikusuma, 2015). Actually, the government has done training aggressively guided by competent persons but in reality it is not an efficient training. The reason is because there is no valid example of an instrument that allows teachers to implement character-based authentic assessment. Assessment in the character dimension is not only the cognitive achievement of learners but also the achievement of affective and psychomotor.

The development of character-based authentic assessment includes two competencies they are attitude competences and skill competences done in a balanced, so it can be used to determine the relative position of each learner against established standards. In Permendikbud No. 66, 2013 stated that the scope of assessment refers to the scope of the material, the competence of subjects, the competence of the charge/competency programs and processes.

Character-based authentic assessment instruments are indispensable for teachers in assessing the achievement of student characters. But unfortunately there is no example of a character-based authentic assessment instruments which can be used as a guidance in the implementation of assessment in primary schools. Therefore, in this research, the researcher will introduce and develop character-based authentic assessment instruments as research products. With the authentic character-based assessment instruments it is expected to facilitate teachers in conducting an assessment of character especially in the realm of attitudes and skills in the classroom and outside the class-room.

MATERIALS AND METHODS

Authentic assessment: Authentic assessment refers to assessment tasks that resemble reading and writing in the real world and in school. Its aim is to assess many different kinds of literacy abilities in contexts that closely resemble actual situations in which those abilities are used. For example, authentic assessments ask students to read real texts to write for authentic purposes about meaningful topics and to participate in authentic literacy tasks such as discussing books, keeping journals, writing letters and revising a piece of writing until it researches for the reader. Both the material and the assessment tasks look as natural as possible. Furthermore, authentic assessment values the thinking behind work, the process as much as the finished product.

Performance assessment is a term that is commonly used in place of or with authentic assessment. Performance assessment requires students to demonstrate their knowledge, skills and strategies by creating a response or a product. Rather than choosing from several multiple-choice options, students might demonstrate their literacy abilities by conducting research and writing a report, developing a character analysis, debating a character's motives, creating a mobile of important information they learned, dramatizing a favorite story, drawing and writing about a story or reading aloud a personally meaningful section of a story. For example, after completing a first-grade theme on families in which students learned about being part of a family and about the structure and sequence of stories, students might illustrate and write their own flap stories with several parts, telling a story about how a family member or friend helped them when they were feeling sad.

Working on authentic tasks is a useful, engaging activity in itself it becomes an "episode of learning" for the student. From the teacher's perspective, teaching to such tasks guarantees that we are concentrating on worthwhile skills and strategies. Students are learning and practicing how to apply important knowledge and skills for authentic purposes. They should not simply recall information or circle isolated vowel sounds in words they should apply what they know to new tasks. For example, consider the difference between asking students to identify all the metaphors in a story and asking them to discuss why the author used particular metaphors and what effect they had on the story. In the latter case, students must put their knowledge and skills to research just as they might do naturally in or out of school. Hosnan claims that judgment or authentic assessment is statistically significant assessment on the study of students to the realm of attitudes, skills and knowledge. In academic life, phrases of authentic assessment and authentic assessment are often interchangeable. Muslich (2009) states, the assessment of the actual (authentic assessment) is the process of collecting a variety data that can give an idea or information about the development of student's learning experiences. According to Sikumbang (2014) various types of authentic assessment are performance assessment observations and questions, presentation and discussion, projects and investigations, portfolios and journals. The authentic assessment is to assess the knowledge and skills (performance) which is obtained by students. Appraiser is not only teachers but it can also be a friend of another or others. Character assessment is done through observation, self-assessment, peer assessment and teacher journals. Skill assessment is done

through the performance appraisal, the assessments requiring students to demonstrate a certain competence by using the practice test, project and portfolio assessment. The instrument used in this activity is a checked list or the assessment scale (rating scale) equipped by rubrics.

Government (Kemdiknas/Kemdikbud) has set a standard of educational assessment which can be used as a guidance by the teacher in the assessment of the school, namely Ministerial Regulation No. 20 of 2007 on standards for educational assessment. In this standard, a lot of techniques and forms of assessment offered to conduct the assessment including the assessment of character. In the assessment of the character, the teacher should make an assessment instrument equipped by an assessment rubric to avoid subjective judgment, either in the form of observational assessment instrument (observation sheet) as well as the attitude scale assessment instruments (e.g., Likert scale).

Character values taken by schools as a main value are quoted from the points of competency standards graduation and some subjects in school which are targeted to internalized by learners. Said that the values are as follows:

- Religious
- To be honest
- Tolerance
- Discipline
- Work hard
- Creative
- Independent
- Democratic
- Curiosity
- Having spirit of nationality
- Love for the motherland
- Appreciating the achievements
- Friendly/communicative
- Love peace
- Joy of reading
- Care for the environment
- Social care
- Responsibility

The formats for performance assessments range from relatively short answers to long-term projects that require students to present or demonstrate their research. These performances often require students to engage in higher-order thinking and to integrate many language arts skills. Consequently, some performance assessments are longer and more complex than more traditional

assessments. Within a complete assessment system, however there should be a balance of longer performance assessments and shorter ones.

Some modification appearing on authentic assessments:

We can see some modification appearing on authentic assessments are as follows:

- Involving students on tasks which are important, interesting, useful and relevant to real life of students
- It is as a learning activity, not as a traditional test
- Involving higher level thinking skills which covers a broad knowledge
- Trying to give understanding to students about what to do
- An assessment tool with standard background and not standardized assessment tool
- Student-centered not teacher-centered
- It can assess students with different abilities, learning styles and cultural backgrounds

RESULTS AND DISCUSSION

Location of the research: This research is conducted in SD Negeri 104202 Bandar Setia and SD Negeri 106811 Bandar Setia which is located on Jalan. Terusan Dusun V Bandar Setia, Kecamatan Percut Sei Tuan, Kabupaten Deli Serdang. The research is done from April to December 2016 for the first year and will be taken up again in the second year. Subjects of this study are:

- Teacher
- Students as subjects for testing a limited group

Experts for validation of assessment instruments as follows:

- Mathematician SD
- Expert in IPA SD
- Expert in IPS SD
- Expert in Indonesian language SD
- Expert in PPkn SD
- Expert in giving marks
- Psychologists

The method used in this research is a research method development with a formative research type (Tessmer, 1998). According to Tessmer, this research is focused on the development of two phases: preliminary and formative evaluation that includes self evaluation,

Table 1: Result of expert validation

Component	Sub-component	Score	%	Criteria
Objectivity	The ability of the instrument to measure student's ability	26	92.86	Very good
	The ability of the instrument to measure the students according to the actual circumstances	24	85.71	Good
	Mean		89.29	Good
Systematic	Suitability instrument with learning indicators	19	67.86	Poor
	Assessment instruments arranged in a systematic and coherent	26	92.86	Very good
	Mean		80.36	Good
Construction	Correctness on sentence arrangements on assessment instruments developed	19	67.86	Poor
	Writing of sentence arrangements in assessment instruments developed does not create double meaning	27	96.43	Very good
	Mean		82.14	Good
Language	The use of Indonesia language rule is correct on the writing instrument developed	18	64.29	Poor
	Mean		64.29	Poor
Practicality	It is easy to implement the assessment instruments developed	27	96.43	Very good
	it is easy to use assessment instruments developed in general	24	85.71	Good
	Mean		91.07	Very good
Total			83.33	Good

prototyping (expert reviews, one-to-one and small group) as well as field test. The instruments used in this study are sheet of expert validation and sheets attitude assessment instruments. Instrument validation of character-based authentic assessment questionnaire using Likert scale which is modified by researcher to the scale of 4. Sudjana (2007) suggests that the evaluation criteria are classified in four levels with the following ratings:

- Poor
- Enough
- Good
- Very good

Analysis data is used for validation of character-based authentic assessment instruments developed. To analyze the results of the assessment given by experts and teachers to see the quality and feasibility of the products is used descriptive analysis by using formula of Mean Percentage Score (PRS) expressed by Sudjana (2007) and the internal reliability of all instruments proposed by Guilford in Suherman.

The development of the character based on character assessment: The development of assessment instruments to the character-based attitude is done to make the completeness of the assessment which does not exist at a teacher's book and student's books. Instrument development, the assessment is only done on the whole theme is contained in the fourth grade primary school. The development instrument in character assessment is also based on the types of assessment contained in Curriculum 2013 including observation ratings, self, peers and teachers journal. The followings are forms of assessment instruments developed attitudes, along with the scope of learning, assessment rubrics and scoring information. Based on the acquisition of validation score

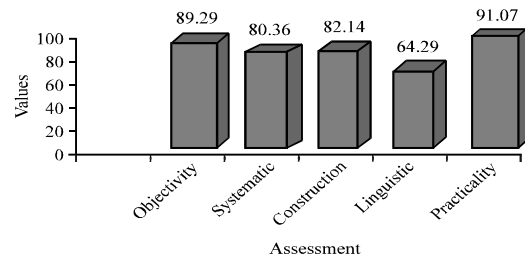


Fig. 1: The assessment result of expert validation

from experts shown in Table 1. The obtaining of mean in each component of character assessment developed based on expert validation is presented in Fig. 1.

Based on the chart, it can be concluded that each component has varied categories, i.e., good and excellent. Overall the average value of 5 components and 9 sub-components components are obtained an average score of 83.33 or has a good category. So, it can be informed based on validator that the validation results to an assessment instrument of character is good and it is feasible to be used to assess student's characters with a note that it should be revised at the sub-components categorized not good.

Field trial: Character assessment instruments revised based on suggestions for improvement of the validator should be validated in field trials. The aim of field trials is to fulfill the elements of the feasibility of assessment instruments attitude at fourth grade elementary school students. Implementation of field trials is divided into two scales, i.e., a small scale test and a large-scale test conducted on students and teachers at fourth grade elementary school in SD Negeri 104202 Bandar Setia and SD Negeri 106.811 Bandar Setia.

A small-scale trial is conducted by one theme: the Beauty of Togetherness and has three sub-themes, namely: the diversity of my national culture; unity in

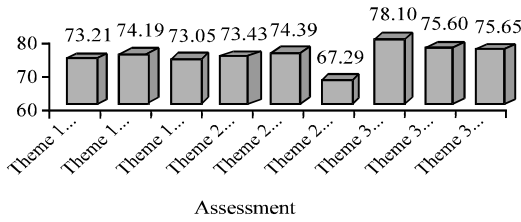


Fig. 2: Score of student character assessment

Table 2: Total score assessment for each theme

Theme to	Sub theme to	Mean score
1	1	73.21
	2	74.19
	3	73.05
2	1	73.43
	2	74.39
	3	67.29
3	1	78.10
	2	75.60
	3	75.65
Total		73.88

Diversity and be grateful for diversity. This a small scale trial is carried out at the first 3 weeks at the school in academic year 2016/2017, precisely on 18th July to 6th August, 2016 by applying the first sub-themes on a weekly basis.

Based on a small scale trial it shows that teachers or students do not have any obstacles in using the instrument at the assessment based on the character of one theme that is the beauty of togetherness. However, students obtain a lower score at the first sub-theme at a peer assessment. It happens because the students give a score for their characters which do not based on reality but based on a subjective or personal elements among students. At the second sub-theme, teachers help researchers to guide students to provide peer assessment objectively in accordance with the real character of his friend and not because of subjective or personal relationships among students. So that the second sub-theme and the third sub-theme students are motivated to achieve character competency which has been determined at each meeting. So at the second sub-theme and the third sub-theme students are motivated to obtain a high score. The trial results of a small scale at the second and third sub-theme have a high score. It indicates that a revision or improvement of assessment instruments is not needed but it only needs a little correction on instructions for the use of instrument.

After that a large-scale test is conducted at the fourth grade with two themes, namely always save energy and care for living things. Each theme has been separated into three sub-themes. Each sub-theme has the idea of always save energy, namely a variety of energy sources the use of energy and force and motion. Sub-theme on the theme of care for living things, namely animal and plants

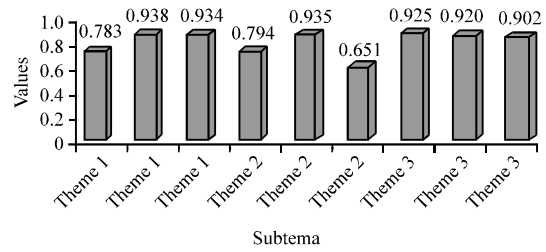


Fig. 3: The result of reliability test

in my neighborhood the diversity of living things in the neighborhood and come on love the surroundings. Large-scale trial is conducted at 4 weeks until 9 weeks in the academic year 2016/2017, precisely on 8 August to 17 September, 2016 by applying one sub-theme for each week.

The results of a large-scale trial shows that teachers or students do not have any problems in using the instrument at the assessment based on the character at the second theme (always save energy) and on the third theme (care for living things). Even, the students obtain a high score for characters for each sub-theme. It happens because students begin to understand that learning in school is not just a cognitive-based but also attitude. Attitudes also affect the value of report so that they do not longer behave something unreasonable. Thus, the results of a large-scale trial shows that it does not need to make be a revision of the assessment instruments as well as the guidance to use the instruments.

The result of field trials shows that the character-based assessment instruments developed attitudes have a good validity. It can be proved from the test attitudes towards student assessment instruments using the same four types of assessments as well (observation assessment, self-assessment, peer assessment and teacher journals votes) generating synchronous data as presented in Table 2.

Based on Table 2 and Fig. 2, it is obtained that the instrument of character assessment on first theme and first subtheme till the assessment on the third theme till third subtheme have an average score of adjacent. Even at total average, the score is not far.

The result of field trials also shows that overall character-based assessment instruments developed attitudes have a very high reliability categories as presented in Table 3.

The result of reliability test is presented at Fig. 3. Based on the expert validation and field trials, it can be concluded that the character-based assessment instrument in elementary school that developed in this research is valid criteria and has a very high reliability. So, it can be used as a measure of student's characters. Expert validation results show that the overall assessment instrument developed is valid or feasible to be used but

Table 3: Instrument reliability on each sub-theme

Theme to	Sub theme to	rb	rii	Category
1	1	0.643	0.783	High
	2	0.884	0.938	Very high
	3	0.876	0.934	Very high
2	1	0.658	0.794	High
	2	0.878	0.935	Very high
	3	0.483	0.651	High
3	1	0.861	0.925	Very high
	2	0.852	0.920	Very high
	3	0.821	0.902	Very high
Total mean		0.770	0.872	Very high

there is a little note to be improved in accordance with the advice of validator. After that it must be there is a test subject or field testing. Based on the results of a small scale and a large scale on each sub-theme, the assessment of teachers, friends or self-assessment on each of the students are not much different. The fact that there are many students having the same score, so the assessment instruments developed an overall declared valid and it can be used as an instrument for student's character assessment. It is based on the average score of the students' character obtained from the assessment on one the first theme and the first subtheme till the third theme and the third subtheme which have an average score of adjacent. Even, at total score also has a score which is not far. The results of field trials also shows that overall character-based assessment instruments developed characters have a very high reliability category.

The findings in this study are supported by several previous studies which have been published in national journals, including the research done by Ramirez *et al.* (2012) indicating that all Indonesian language teachers have implemented an authentic assessment of learning. Teachers apply various models of authentic assessment indicating that there are factors of creativity in the application of assessment. However, teachers are still experiencing various kinds of constraints in implementing authentic assessment. Another research conducted by Abidin (2012) shows that authentic assessment guides learning through various creation of learning activities of the student during the learning process containing of characters. Authentic assessment gives a real picture of student's abilities in reading and provides a measure of the achievement of student's character development. Based on the fact that the use of authentic assessment will contribute to an increase reading comprehension and character development of students.

This research is not only supported by studies published by national journals but also internationally as research conducted by Ramirez *et al.* (2012) shows that the characteristics of students, achievement and attitudes have a statistical value related to one another. Another

research is conducted by Nolan *et al.* (2012) the survey results show a good attitude assessment instruments which can contribute to the measurement of student's attitudes. The results of this study shows that the use of authentic assessment is an innovation learning development that can serve as the vehicle of implementation and attitude measurement through the democratic way. An authentic assessment is able to train teachers to be a model of character development in the learning process and also able to develop student's character.

Based on the findings in this study and several previous studies it can be concluded that the assessment character-based at elementary school developed in this research is valid criteria and has a high reliability, so it can be used as a measurement of student's characters. It can be summarized that the development of an authentic assessment instruments in the realm of character-based attitude to the fourth grade of elementary school is ready to be used and can provide accurate assessment attitude.

CONCLUSION

The results of preliminary observations concerning the existence of an assessment instrument indicates that the character-based authentic assessment instruments have not been applied in elementary school. Expert validation results show that attitude character-based assessment instruments are valid, having a good category and can be used to assess student's attitude but there is a little correction should be done at the sub-components which are categorized not good. The results of field trials shows that the character-based assessment instruments developed attitudes have good validity and high reliability. The evidence can be seen from the trial assessment of the students using the same four types of assessments as well (observation assessment, self assessment, peer assessment and teacher journals) generates synchronous data and the range of scores which are almost the same. Thus, based on the results of expert validation and field trials it can be concluded that the attitude of character-based assessment instruments are valid and have very high reliability category, so the instrument is eligible to be used in measuring the achievement of the values of student's characters.

Here, the researcher would like to give some advices. The teacher should be able to adjust the indicator used for student's characters with material studied, particularly the implementation of curriculum 2013. For other researchers who will conduct research and development, it is better if the product developed is suitable with the subject required as a sample of research and the research

should be done at school having enough facilities, media and participants which can support the research. If other researchers conducting research with the model and the same instrument but different research samples, the result tends not much different from the results of this study and expected to obtain the perfect results.

REFERENCES

- Abidin, Y., 2012. [In the model of authentic assessment of learning in reading comprehension Beroreintasi character education (In Indonesian)]. *J. Character Educ.*, 2: 164-178.
- Hadikusuma, Z., 2015. Authentic assessment based curriculum development in class IV 2013 basic state school city Pekanbaru. *J. Thematic*, Vol. 5,
- Majid, 2014. [Integrated Thematic Learning]. Remaja Rosdakarya, Bandung, Indonesia, (In Indonesian).
- Muslich, M., 2009. [Implementar PTK]. Bumi Aksara Publisher, East Jakarta, Indonesia, (In Indonesian).
- Nolan, M.M., T. Beran and K.G. Hecker, 2012. Surveys assessing students attitudes toward statistics: A systematic review of validity and reliability. *Stat. Educ. Res. J.*, 11: 103-123.
- Ramirez, C., C. Schau and E. Emmioglu, 2012. The importance of attitudes in statistics education. *Stat. Educ. Res. J.*, 11: 57-71.
- Sikumbang, R., 2014. [Scientific Approach and Contextual Learning 21st Century: The Key to Successful Implementation of the Curriculum in 2013]. PT. Ghalia Indonesia Printing Publisher, Bogor, Indonesia, ISBN:9789794506837, Pages: 454 (In Indonesian).
- Sudjana, 2007. [Methods Statistics]. Tarsito Publisher, Bandung, Indonesia, (In Indonesian).
- Tessmer, M., 1998. Planning and Conducting Formative Evaluation. Kogan Page Publisher, London, England.