The Validity of Text Based Evaluation Learning Instruments: Reconstruction of Character Evaluation Instrument Models in the Indonesian Language Learning

Syaiful Musaddat*
Language Education Program
Faculty of Teacher Training and Education
Universitas Mataram
NTB, Indonesia
syaiful_musaddat@unram.ac.id

Imam Suryadi
Language Education Program
Faculty of Teacher Training and Education
Universitas Mataram
NTB, Indonesia
imamsuryadi@unram.ac.id

Siti Rohana Hariana I.
Language Education Program
Faculty of Teacher Training and Education
Universitas Mataram
NTB, Indonesia
rohanahariana@unram.ac.id

Suyanu Suyanu
Language Education Program
Faculty of Teacher Training and Education
Universitas Mataram
NTB, Indonesia
suyanu@unram.ac.id

Abstract – This study aims to describe the validity of text-based language learning evaluation instruments developed by the Alumni of Indonesian Language Education Program of Mataram University who teach at Senior High Schools throughout West Lombok Regency in order to reconstruct the character evaluation instrument models in the Indonesian language learning. The aspects studied include construct validity (choice of tools), content validity (completeness of the tools) of evaluation developed, and character evaluation instrument models that can be used in learning Indonesian language. Data collection was done through the document review method and Focus Group Discussion (FGD). The data were gotten from learning tools that developed by the Alumni who teach in high schools throughout West Lombok Regency. In this case, the Alumni who have at least 5 years teaching experience and have been certified as professional educators. Data analysis was carried out by following the principles of qualitative research, i.e. data reduction, data presentation, and data verification or interpretation. Based on the results and discussion, it can be concluded that: (1) The construct validity are (a) the recognition context and modeling stage: 60% high, 40% medium, and 0% low; (b) stages of building texts in groups: 50% high, 40% medium, and 10% low, and (c) stages of constructing individual texts: 70% high, 30% medium, and 0% low. (2) The content validity are (a) the recognition context and modeling stage: 50% high, 40% medium, and 10% low; (b) stages of building texts in groups: 40% high, 50% medium, and 10% low, and (c) stages of constructing individual texts: 60% high, 30% medium, and 10% low. (3) The character evaluation instrument model that can be developed in language learning is a process evaluation model in the form of an attitude or behavior observation sheet at all stages of learning. Thus, it is advisable for the Indonesian Language Education Program at Mataram University to immediately take some concrete steps in order to improve the quality of its alumni, especially in developing the learning evaluation instruments.

Keywords – construct validity, content validity, and character evaluation instrument

I. INTRODUCTION

The results of preliminary observations made through the Teacher Professional Education and Exercise program, Lecturer Assignments in Schools, and Teacher Professional Education shows that the quality of Indonesian language learning at schools have not been optimal nowadays. Things that prove this fact are as follows. First, in terms of lesson planning: (a) teachers are often found teaching without lesson plan; and (b) the lesson plans are not detailed and are rarely equipped with students' worksheets and assessment instruments. Second, in terms of the learning implementation: (a) teaching and learning process often done without planning; and (b) poor learning reinforcement, both praise of success and the improvement of failure. Third, in terms of learning evaluation: (a) assessment of learning outcomes tends to measure cognitive/knowledge aspects; (b) evaluations tend to see the final result (product evaluation) without evaluating the process; and (c) the unavailability of adequate learning evaluation tools.

The results of training through the community service done by Musaddat, et al. [1] and Intiana, et al. [2] show the need for teachers' guidance (including alumni). In this case, it is necessary to foster teacher's ability to develop learning tools (lesson plan, student worksheet, teaching materials, and evaluation instruments). The study of Musaddat [3] shows that the quality of Indonesian language teachers (Alumni of Indonesian Language Education Program) in planning and implementing learning in general (without including text-based learning) was in the good category. In planning the lesson plan, the category was good with average 4.0, learning implementation was also categorized good with mean 4.2. However, these facts ware far from the expectation.

The application of the 2013 curriculum requires that learning Indonesian must use a text-based learning model. In this case, the planning, implementation, and evaluation of learning must follow the demands of text-based learning. Musaddat's study result [4] shows that the competence of Indonesian language teachers (Alumni) in West Lombok in designing text-based learning is still not optimal. It was found that the developed lesson plans did not draw the stages of the 2013 curriculum for the text-based learning. At the implementation stage of learning, the stages of text-based learning also did not run optimally, there were stages sometimes missed by the teacher [4] and students [5].

To develop a text-based learning model according to 2013 curriculum demands, comprehensive data are needed related to planning, implementing, and evaluating text-based learning. Data for planning and implementing learning has been obtained through previous studies in 2016 and 2018. Meanwhile, data related to evaluation of text-based learning has not been obtained. Therefore, it is necessary to study the evaluation of text-based learning through research.

Based on the issue above, this paper was aimed at discussing the validity of the text-based learning evaluation instrument at the stages of: (1) building contexts and modeling, (2) developing texts in groups; and (3) developing texts individually. Further, this will be used as material in reconstructing the character evaluation model or social behavior within text based learning.

Related to the implementation of the 2013 curriculum, an evaluation of learning that should be done is a class-based assessment that is comprehensive and sustainable. The curriculum center in Muslich, [6] explains that classroom-based assessment is an activity of gathering information about the process and student learning outcomes conducted by the teacher in order to see what students want to measure. It is not only the students; final ability acquired in this context, but also the development of student learning experiences. This is identical with the authentic assessment in a contextual approach. According to this view, assessment is directed at the process of observing, analyzing, and interpreting data that has been collected in the learning process of students, not merely on the learning outcomes [7].

Related to the above exposition, assessment must be planned before, during and after the learning process. Another thing that must be planned is the choice of evaluation tools and techniques. In terms of technique, the teacher can use a test or non-test. If a test is used, it must be well planned. Likewise, if a non-test is used, it must be chosen and planned well so that it is truly on target. In connection with the test technique, there are two types of learning outcomes tests that are used to measure students' abilities, which are standard tests and teacher-made tests [8]. Standard tests are tests that have been standardized through a standardization process, so they have an adequate level of validity and reliability [9]. Teacher-made tests or informal tests are tests that teachers make and use in the classroom learning activities. Teacher-made tests are tests that are prepared, carried out, and scored by a teacher to determine the achievement of student competencies.

The development of an evaluation instrument is a procedure for making a process evaluation instrument and a rubric for evaluation results. Indonesian language learning which emphasizes more on the aspects of skills/ skills requires assessment instruments that are focused on performance, such as observation sheets and assessment rubrics. Procedures or steps in preparing a learning assessment is in the form of observation sheets and performance appraisal rubrics that can be guided by the stages that have been developed by Nurgiantoro [10] and Musaddat [11].

Text-based learning is a learning process that makes text as the basis, principle, and foundation [12]. Learning methods in the text-based learning are carried out with three stages in a cycle, i.e.: (1) introduction of context and modeling, (2) joint text development, and (3) independent text development [13] [14]. In this case, text-based learning is taught through three categories of learning activities, i.e.: context recognition and modeling, joint text development, and independent text development [15], [16], and [14]. Therefore, the instruments needed for text based learning must refer to these stages. It means that assessment instruments are needed at all three stages where the required instruments are in the form of process evaluation instruments and product/ results assessment instruments.

II. METHOD

This research was carried out through three stages: (1) the pre-research or proposal preparation stage, the preparation of research instruments, and the preparation of research tools and materials; (2) data collection and analysis; and (3) post-research stage which includes the preparation of research reports. Data collection was done through the document review method and Focus Group Discussion. The data were taken from the learning tools developed by the alumni of Indonesian Language Education Program of Mataram University who teach in high schools in West Lombok and the FGD results.

III. FINDINGS AND DISCUSSION

This section describes the results of research and discussion, related to the construct validity and the contents validity of the text-based assessment instrument developed by the alumni of Indonesian Language Education Program, Mataram University. The presentation adjusted to the formulation of the problem discussed which is successively according to the stages of text-based learning. In this case, the construct validity and the content validity of the text-based assessment instrument are at: (1) building context and modeling, (2) developing text in groups; and (3) developing texts individually. Then, a character evaluation model was reconstructed in text-based learning.

A. The construct validity of the Text-Based Learning Evaluation Instrument

Based on the tabulation and results on the analysis data, a number of evaluation model of Bina Sarana Informatika (BSI) learning were found. From this evaluation model, it is found that there are high, medium, and low validity evaluation models. The determination level of validity was done by comparing the learning indicators with the type of

evaluation instrument used [9]. It is said to have high validity if all evaluation tools developed are in accordance with competency indicators in the curriculum. Validity is said to be medium if most of the evaluation tools developed are in accordance with competency indicators in the curriculum. Meanwhile, validity is said to be low if only a small number of evaluation tools are developed in accordance with competency indicators in the curriculum [18].

1) Introduction Context and Modeling Stage

Based on text-based learning guidelines, the recognition of context and modeling has two main indicators that students should reach. The two indicators are meant to understand the structure of the task and understand the language of the text. In the context of evaluating the learning outcomes, this is a measure of the construct validity.

In the lesson plan produced by all respondents, two forms of evaluation instruments used were found, namely the description task and multiple choice questions. The results analysis of the assessment instruments found that 60% were categorized as high, 40% were medium, and 0% is low. In this context, there are six instruments in which all evaluation tools are developed in accordance with competency indicators in the curriculum. Furthermore, there are four instruments, most of which are evaluation tools that are developed in accordance with competency indicators in the curriculum and there are no instruments with low validity at this stage.

2) Stage of Building Text in Groups

Based on the text-based learning guidelines, at the text development stage in groups, there is one main indicator that students must achieve, namely the ablility to develop texts together. In the context of evaluating learning outcomes, of course this is a measure of the validity of the construct.

In the lesson plan developed by all respondents, a form of evaluation instrument was found and it is description task in the form of performance order (compiling the text together). The result analysis shows that 50% of the assessment instruments was categorized high, 40% was medium, and 10% was low. Hence, there are five instruments in which all evaluation tools are developed in accordance with competency indicators in the curriculum. Furthermore, there are four instruments, most of which are evaluation tools developed in accordance with competency indicators in the curriculum and one instrument with low validity where only a small number of evaluation tools that are developed in accordance with competency indicators in the curriculum.

3) Stage of Building Individual Text

Based on the text-based learning guidelines, at the individual text development stage there is one main indicator that students must reach, namely the ability to develop texts independently. In the context of evaluating learning outcomes, it was found only one evaluation instrument used in the lesson plan of all respondents, namely description task in the form of performance order (compile the text independently). The results of the analysis of the assessment instruments showed that 70% were in high category, 30% were in medium category, and 0% was in low category. There are seven instruments in which all evaluation tools are developed in accordance with curriculum competency indicators. Furthermore, there are three instruments where most evaluation tools are developed in accordance with

competency indicators in the curriculum and no instruments with low validity are found.

Meanwhile, there were four types of evaluation process found, namely: observation sheets, self-assessment, peer assessment, and teacher's daily journal. However, not all respondents' lesson plans are equipped with the assessment process instrument. In terms of its construct, the average assessment instruments process have high validity. From the lesson plan explanation, it is found that this process of evaluation instrument is used at all stages of learning.

B. The Content Validity of the Text Based Learning Evaluation Instrument

The findings of construct validity are then measured in terms of the content validity. In this case, the benchmark is the completeness of each evaluation instrument in terms of learning indicators, whether or not there are guidelines for scoring, and whether or not the assessment rubric is complete [9]. It is said to have high validity if all the learning indicators are accepted by the instrument, equipped with guidelines for scoring and assessment rubrics. Medium validity comes when one of the three requirements (completeness of contents, scoring guidelines, and rubrics) are not fulfilled. Meanwhile, it is said that the validity is low if it only fulfills one of the three requirements (completeness of contents, scoring guidelines, and rubric) [18].

1) Introduction Context and Modeling Stage

There are two forms of outcome evaluation instruments were found in this stage, namely: description task and the multiple choice questions. The analysis results of the content validity on the assessment instruments revealed that 50% were high, 40% were medium, and 10% were low. In this context, there are five instruments that fulfilled the three requirements of content validity, they are content completeness, scoring guidelines, and rubrics. Furthermore, there are four instruments fulfilling two requirements, and one instrument fulfilling one requirement.

2) Stage of Building Text in Groups

In accordance with the analysis results on the point of construct validity analysis, there is a form of evaluation instrument of the results used, that is a matter of description task in the form of a performance order (compiling the text together). The results of the analysis content validity of the assessment instruments revealed that 40% were in high category, 50% were in medium category, and 10% were in low category. In this context, there are four instruments that meet the three requirements for content validity, namely: completeness of content, scoring guidelines, and rubrics. Furthermore, there are five instruments fulfilling two requirements, and one instrument fulfilling one requirement.

3) Stage of Building Individual Text

In accordance with the results of the analysis on the point of construct validity analysis, it was found one form of evaluation instrument of the results used, namely description task in the form of a performance order (compiling the text independently). The results analysis of the content validity on the assessment instruments revealed that 60% were in high, 30% were in medium, and 10% were in low category. There are six instruments that meet the three requirements for content validity, i.e. completeness of content, scoring guidelines, and rubrics. Furthermore, there are three

instruments fulfilling two requirements, and one instrument fulfilling one requirement.

From the four types of process for instruction evaluation (observation sheets, self-assessments, peer evaluations, and teacher's daily journals) found, the average process evaluation instrument was on the medium category. Many of the assessment rubrics (descriptors or guidelines) are not specific, less appropriate, and some are not appropriate at all. For this matter, further study is needed due to the lack data collection.

C. Character Evaluation Instrument Models and Social Attitudes of Text Based Learning

Based on the findings of the results and discussion, one of the character evaluation instrument models that can be developed in language learning is the process assessment model in the form of an attitude observation sheet or social behavior at all stages of learning. This instrument must be valid in terms of form (the aspect of observation must be in accordance with the character / social attitude in the Basic Religious Competition) and valid in terms of content (equipped with scoring guidelines and rubrics/clear description, specific, and decisive). In this case, the character of evaluation instrument model that can be developed in language learning is a process evaluation model in the form of an attitude or behavior observation sheet at all stages of learning/performance evaluation model [19]. Attitude assessment also has a high level of validity in learning [20]. The following is a model that can be used as an instrument for character assessment/social attitude.

TABLE I.   MODEL CHARACTER ASSESSMENT INSTRUMENT / SOCIAL ATTITUDES

No	Name	Assessment Aspect *				Acquisition Aspect	Score
		Honest	Discipline	Responsible	Socialist		
1.							
2.							
3.							
Etc							

* Fill in the score according to the following descriptor.

TABLE II.   RUBRIC ASSESSMENT OF CHARACTER ASPECTS AND SOCIAL ATTITUDES

ASSESMENT ASPECT	DESCRIPTOR AND SCORE			
	VERY GOOD (10)	GOOD (7)	LESS GOOD (5)	NOT GOOD (1)
Honest	Always appropriate between words and actions at each stage of learning	Occasionally inappropriate between words and actions at each stage of learning	Often inappropriate between words and actions at each stage of learning	Always inappropriate between words and actions at every stage of learning
Discipline	Always on time in completing the assignments or following the learning stages	Often on time in completing the assignments or following the learning stages	Rarely on time in completing the assignments or following the learning stages	Never on time in completing the assignments or following the learning stages
Responsible	All tasks that becomes responsibility can be completed	Most of the tasks that becomes responsibility can be completed	Fraction of tasks that becomes responsibility can be completed	All tasks that becomes responsibility cannot be completed
Socialize	Can work with all members of the group	Can work with most members of the group	Can work with a small group of members	Cannot cooperate with all group members

IV. CONCLUSION AND SUGGESTION

Based on the findings and discussions above, it can be concluded that: (1) the construct validity are (a) the stage of context recognition and modeling, in which 60% were in high, 40% were in medium, and 0% was in low category; (b) stages of building texts in groups, in which 50% were in high, 40% were in medium, and 10% were in low category, and (c) stages of constructing individual texts, in which 70% were in high, 30% were in medium, and 0% were low category. (2) The content validity are (a) stage of context recognition and modeling, in which 50% were in high, 40% were in medium, and 10% were in low category; (b) stages of building texts in groups, in which 40% were in high, 50% were in medium, and 10% were in low, and (c) stages of constructing individual texts, in which 60% were in high, 30% were in medium, and 10% were in low category. (3) The character evaluation instrument model that can be developed in language learning is a process evaluation model in the form of an attitude or behavior observation sheet at all stages of learning.

Thus, it is recommended to the Indonesian Language Education Program at Mataram University to immediately take concrete ways in order to improve the quality of its alumni, especially in developing the learning evaluation instruments.

REFERENCES

[1]   Musaddat, S. et.al. 2018. *Pelatihan Pengembangan Materi Ajar Berbasis Teks Bagi Guru-guru Bahasa Indonesia Aluni PBSI FKIP Unram di Kabupaten Lombok Barat,* Jurnal Pendidikan dan Pengabdian Masyarakat Vol. 1 No.2, Agustus 2018 (page. 321–326)

[2] Intiana, Siti R.H. et.al. 2018. *Pelatihan Pengembangan Materi Ajar Bagi Guru-guru Bahasa dan Sastra Indonesia di Kabupaten Lombok Utara,* Jurnal Pendidikan dan Pengabdian Masyarakat Vol. 1 No1, Februari 2018 (page. 165–173)

[3] Musaddat, S. et.al. 2016. "Pemetaan Kualitas Alumni Program Studi Pendidikan Bahasa, Sastra Indonesia, dan Daerah (PBSID) FKIP Universitas Mataram dalam Hal Perencanaan dan Pelaksanaan Pembelajaran (Laporan Penelitian)." Mataram: LPPM Unram

[4] Musaddat, S. et.al. 2019. *Pemetaan Kompetensi Guru Bahasa Indonesia PBSI FKIP Unram dalam Merancang dan Melaksanakan Pembelajaran Berbasis Teks,* Jurnal Mabasindo Vol. 3 No. 1 Edisi Mei 2019 (page. 136–148)

[5] Dharma, Purnama S.V. et.al. 2019. *Implementasi Pembelajaran Bahasa Indonesia Berbasis Teks di SMA Negeri 1 Bengkulu Tengah Kelas XI,* Jurnal Ilmiah Korpus, Vol. III, No. I, April 2019 *(page. 66–74)*

[6] Muslich, Masnur. 2008. *Pembelajaran Berbasis Kompetensi dan Kontekstual [Competency- and contextual-based learning*]. Jakarta: Bumi Aksara

[7] Nurhadi dan Agus Gerrad Senduk. 2004. *Pembelajaran Kontekstual dan Penerapannya dalam KBK [Contextual learning and it application in KBK]* Malang: Universitas Negeri Malang.

[8] Sudjana, N. 2000. *Penilaian Hasil Proses Belajar Mengajar [Evaluation of Learning process].* Bandung: Remaja Rosdakarya.

[9] Djiwandono, M. Soenardi. 2008. *Tes Bahasa dalam Pengajaran [Language test in its teaching].* Jakarta: Indeks

[10] Nurgoantoro, B. 2001. *Penilaian dalam Pengajaran Bahasa dan Sastra [Evaluation in language and literature teaching].* Yogyakarta: BPFE

[11] Musaddat, S. 2013. *Pendidikan Bahasa dan Sastra Indonesia Kelas Tinggi* [Language and literature education in higher grades]. Mataram: Argapuji Press.

[12] Sufanti, Main. 2017. *Pembelajaran Bahasa Indonesia Berbasis Teks: Belajar dari Ohio Amerika Serikat,* https://publikasiilmiah. ums.ac.id/bitstream/handle/11617/3363/ Accessed: 13/09/19

[13] Depdikbud. 2017. *Bahasa Indonesia SMA/MA Kelas XII [Indonesian language for SMA grade XII]. Jakarta*: Depdikbud.

[14] Mahsun. 2018. *Pembelajaran Bahasa Indonesia Berbasis Teks (Edisi Kedua) [Text-based Teaching Indonesian language].* Jakarta: Rajawali Press

[15] Depdiknas. 2002. *Model Pembelajaran Bahasa Indonesia [Teaching Indonesian language]. Jakarta*: Dirjen Dikdasmen.

[16] Isodarus P. Baryadi. 2017. *Pembelajaran bahasa Indonesia Berbasis Teks,* Jurnal Ilmiah Kebudayaan SINTESIS, Vol. 11, No. 1, Maret 2017 (page. 1-11)

[17] Mahsun. 2010. *Metode Penelitian Bahasa [Language Research method].* Jakarta: Rajawali Press

[18] Musaddat, S. et.al. 2017. "Pemetaan Kualitas Alumni Program Studi Pendidikan Bahasa, Sastra Indonesia, dan Daerah (PBSID) FKIP Universitas Mataram dalam Hal Mengembangkan Intrumen Evaluasi Pembelajaran (Laporan Penelitian)." Mataram: LPPM Unram

[19] Yuniawan, Tommi. 2014. *Model Penilaian Kinerja dalam Pembelajaran Membaca Berbasis Teks Narasi Bermuatan Pendidikan Karakater Cinta Budaya.* Jurnal Pendidikan Karakter**,** Tahun IV, No. 1, Februari 2014 (page. 61 -72)

[20] Mulyana, Edi H., et.al. 2016. *Pengembangan Penilaian Sikap dengan Menggunakan Analisis Video Pelaksanaan Pembelajaran Berbasis Kontek.* Prosiding Seminar Nasional Sains Tahun 2016 (page 367-373)