ANBUSO: PRACTICAL SOFTWARE TO PERFORM ITEM ANALYSIS

Ali Muhson[1], Barkah Lestari[2], Supriyanto[2], Kiromim Baroroh[2]

[1]Pendidikan Ekonomi FE Universitas Negeri Yogyakarta
[2]Pendidikan Ekonomi FE Universitas Negeri Yogyakarta
[3]Pendidikan Ekonomi FE Universitas Negeri Yogyakarta
[4]Pendidikan Ekonomi FE Universitas Negeri Yogyakarta
alimuchson@yahoo.com

Abstract—*This study develops item-analysis software, and verifies its feasibility. This study uses a model of Research and Development (R & D). The procedure includes designing and developing a product, validating, and testing the product. The data were collected through documentations, questionnaires, and interviews. This study successfully developed item analysis software, namely AnBuso. It was practical and applicable for teachers to analyze test items.*

Keywords: feasibility, item-analysis software, AnBuso

I. INTRODUCTION

Learning processes determine the quality of education. This implies that teachers have a very central role in improving the quality of education. They have crucial roles not only in designing and implementing but also in assessing the learning processes and their outcomes. Improving the quality of learning and assessment system is one of the efforts to improve the quality of education.

Learning and assessments can be said to be 'two sides of the same coin' [1]. All forms of assessment are essential components of teaching [2] and they have strong influences on the quality of learning [3]. Assessments involve not only finding out what students have learned but also improving the processes of learning and teaching [4]. This implies that assessments have essential roles in improving not only the learning processes but also the learning outcomes. Good assessments will encourage teachers to determine appropriate instructional strategies and motivate students to study harder.

An assessment is a process of gathering evidence of student learning to inform instructional decisions [5]. It means that teachers have to gather accurate evidence of the achievements of their students as the quality and impacts of their instructional decisions depend on it. The development of assessment instruments should therefore measure all aspects of learning objectives. Good students should attain better scores than the poor ones.

The implementation of the new curriculum, known as the 2013 curriculum, in the secondary education levels in Indonesia is facing many problems [6], for example the difficulty of changing teachers' mindset [7, 8]. A study reported that 87 percent of Indonesian teachers had difficulties in understanding the implementation of authentic assessments [9] as the assessment system includes too many aspects that must be measured [10]. It means that their ability to perform good assessments has to be improved.

An assessment is a systematic process of providing credible evidence of resources, implementation, and learning outcomes with the purpose of improving and developing effective instruction, programs, and services in higher education [11]. Thus, it is a process of gathering information about the students and in turn the information is employed to make instructional decisions and improve learning processes.

An assessment should be able to assess students' progress. Such information can be used as a basis for identifying the students' status in the group and determining the next steps of learning. Therefore, the assessment has to be able to accurately assess their competence. Assessments help teachers clarify the goals and the purposes of their teaching, create learning experiences in real-life contexts, and provide many different ways for students to demonstrate their abilities and skills [12]. Properly used assessment procedures can contribute directly to the

improvement of learning, such as (1) clarifying the nature of the intended learning outcomes, (2) providing short-term goals to work toward, (3) providing feedback concerning learning progress, (4) providing information for overcoming learning difficulties and for selecting future learning experiences, and (5) identifying the next instructional goal [13]. The procedures involve interrelated tasks in that one step determines the next.

The purposes of assessments include establishing a classroom that supports learning, planning and conducting instruction, placing students, providing feedback, diagnosing student problems and disabilities, and summarizing and grading academic learning and progress [2]. Thus, assessments are used to monitor the learning progress, provide information as a basis for giving feedback, and improve learning.

To provide accurate information, assessments have to be based on several criteria, such as valid, reliable, competence focused, comprehensive, objective, and educative [14]. The three concepts—validity, reliability, and objectivity—have to do with the quality of the information obtained from tests or other assessment instruments or methods [15, 16]. Thus, the validity and reliability are important parts of assessments to obtain appropriate and accurate information.

Item analyses play a somewhat more important role in the construct and predictive validation [17]. It means that the item analysis is an important process, which determines the validity of items and therefore it has an important role in identifying good items. The assessment results also provide accurate information about students' abilities therefore they can help teachers to identify which materials are difficult. The item analysis also provides information about the materials that have not been mastered by the students so that teachers can plan remedial activities. Based on the reasons it is necessary to develop a practical software that can help teachers analyze items practically.

There are several different types of item analysis software available in the market, but they have not been used optimally yet by Indonesian teachers. This is because most of the software use foreign languages which are difficult for them to understand. Some software are also less practical and applicable because they produce output with different formats. Therefore, it is necessary to develop more practical and applicable software that can cater for their needs. This present study aims to develop item analysis software that could be used by Indonesian teachers to prepare learning administration reports. The software is named AnBuso.

II.   METHODS

This study uses a model of Research and Development (R & D). The procedure includes designing and developing a product, validating, and trying the product. A preliminary study was carried out to determine the ability and willingness of teachers to use item analysis and identify their real needs. The next step was designing a model of application software and its guidebook, which were validated by experts and revising them in order to obtain validated products.

The product validation was conducted using the Delphi method. This validation processes also involved some experts in education, measurements and computer programming. The testing of the trial products were conducted twice: limited and extensive testing. The test was to obtain comprehensive information and input from users. The results show that all the instruments, the software and its guidebook are feasible to use although there are some suggestions for revision. The results of this test were also used to improve the quality of the products so that they are ready to be disseminated widely.

The study involved 65 respondents comprising teachers and school supervisors from five districts in Yogyakarta (DIY). Purposive sampling was carried out to select the respondents with computer literacy, especially the ability to operate Microsoft Excel program.

Observation, documentation, questionnaires, and interviews were employed to elicit the research data. Observation was conducted to obtain data or information about the teachers' ability to use the application program that has been developed. The data of instruments used to measure students' abilities, such as formative and summative tests were obtained by documentation techniques. The elicited documents took the forms of final examination tests, daily tests, remedial programs, and the results of the previous item analysis used by the teachers.

Table 1. Software Feasibility Categories

No	Skor	Kategori
1	More than M + 1,8 SD	Very feasible
2	M + 0,6 SD to M + 1,8 SD	Feasible
3	M − 0,6 SD to M + 0,6 SD	Fair
4	M − 1,8 SD to M − 0,6 SD	Less feasible
5	Less than M − 1,8 SD	Not feasible

The questionnaire elicited feedback or suggestion from teachers, supervisors, education department officers, and education experts. The questionnaire was also developed to measure the feasibility of the software concerning its visual display, contents, usefulness, practicality, and convenience. Interviews were conducted to teachers, supervisors, education department officers, and experts to obtain information regarding the strength and drawbacks of the AnBuso software.

The data were analysesed descriptively and supported by tables and graphs. Five scales of categories were employed to measure the feasibility of the software (see Table 1).

III. FINDING AND DISCUSSION

The study involved 65 participants comprising teachers, supervisors and education actors. Most of them (72%) were teachers of high schools (SMA, SMK, and MA). Many of them (68%) were teachers of public schools and 57% were government employees. Respondents came from various fields of study, including Economics, Accounting, Mathematics, Social Studies, English, Indonesian, Information Technology, Chemistry, Physics, Biology, Geography, Arabic, and so on.

The willingness of the teachers to analyze test items was low (57%). Only 11% of them regularly conducted test item analysis and 12% of them never did it. It would be important to assess why the teachers had a low willingness to analyze test items.

In general, the teachers would perform item analysis if the supervisors ordered them to do so. It means that their awareness to perform item analysis was very low. This is because they were unable to use the available software. Of course this contributed to the poor quality of test items they have developed.

This study successfully developed AnBuso software and its manual book. AnBuso was developed on Microsoft Excel program. It has been improved based on the feedback from the respondents of the present research. The AnBuso was accompanied with a manual book in Indonesian language that will help teachers take advantage of the software. It provides step-by-step instructions which are very practical for the teachers to operate the software. The book contains an introduction, contents, input and data report.

The software has undertaken some important modifications in order to be compatible with the assesments system of the 2013 curriculum. Some revisions were conducted to improve its visual display, **Input01** spreadsheet, **Participant Reports** spreadsheet, **Participant Remedial** spreadsheet, and formula changes. The modifications on the visual displays had to be done because it had too many color variations, which have made the application less attractive. The adjustments were done not only on the input spreadsheet, but also on the report spreadsheet (see Figure 1). They were made based on the respondents' feedback.

| Previous visual display | New visual display |

Figure 1. The spreadsheet of Input01 before and after modifications

Figure 2. Grade Revision

AnBuso has been adjusted to assessment system of the 2013 curriculum of the secondary education levels in Indonesia so that it accommodates Indonesian teachers' needs. Some amendments were also made on the assessment scales. In the previous model, the grading scales were 1-10 and 1-100. To cater for the assessments system of the 2013 curriculum, the new model provides a 1-4 grade scale (see Figure 2).

Basically the purpose of the item analysis is to determine the quality of the test items and obtain information about the assessment results so that teachers can make a follow-up plan such as remedial and enrichment teaching. The software also provides information about the students who are included in the remedial and enrichment group.

Figure 3 shows the general overview of the AnBuso system. Basically the software is divided into three components: input, process and output. Data input consists of **Input 01** which takes general data, and **Input 02** which contains the identity of students and their answers. The data entered into the **Input 01** spreadsheet include the School Name, Subject, Class/Program, Name of test, SK/KD, Name of the Teacher, Teacher's NIP, Semester, Academic Year, Test Date, Inspected Date, name of the school Principal, Principal's NIP, Reporting Place, Reporting Date, Grades (10 or 100), the score of Minimum competence criteria (KKM), Objective Test Data, and Essay Test Data. The data entered into the **Input 02** spreadsheet include Students Name, Gender (M = Male, P = Female), Multiple Choice Answer sheet, and Essay Score. Whenever the data are entered in the input 01 and input 02 spreadsheets, the program will automatically generate a complete report consisting of:

1. **Student Report Sheet** which displays list of test scores. The list contains the name of student and gender, number of items answered correctly and incorrectly, objective test score, essay test score, final score, and description (passed and failed), The other information in this sheet is the total number of the students, the number (percentage) of students who passed the tests, the minimum scores, maximum scores, average scores, and the standard deviation.
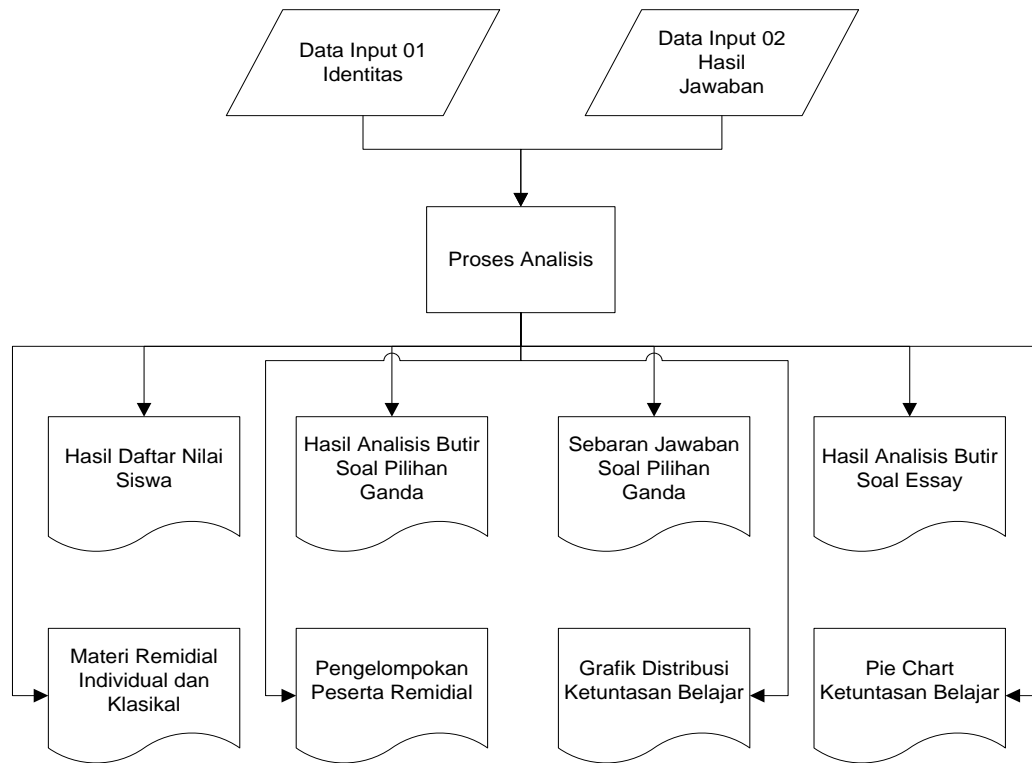
Figure 3. Flowchart Diagram Revision

2. **Item Report Sheet** which shows the result of Multiple Choice Analysis. It contains Discriminant Index, levels of difficulty, Malfunction Distractors, and Conclusion.
3. **Response Pattern Report Sheet** which contains Distribution of the aswers of Multiple Choice tests.
4. **Essay Report Sheet** that contains Discriminant Index, levels of difficulty, and Conclusion.
5. **Remedial Material Reports Sheet**, which contains materials for Individual and Classical Remedial tests.
6. **Remedial Group Reports Sheet**, which contains the grouping of participants who take remedial tests.
7. **Graphics Sheet**, which contains charts of minimum competence criteria (KKM) score distributions, and the proportions of the learning mastery.
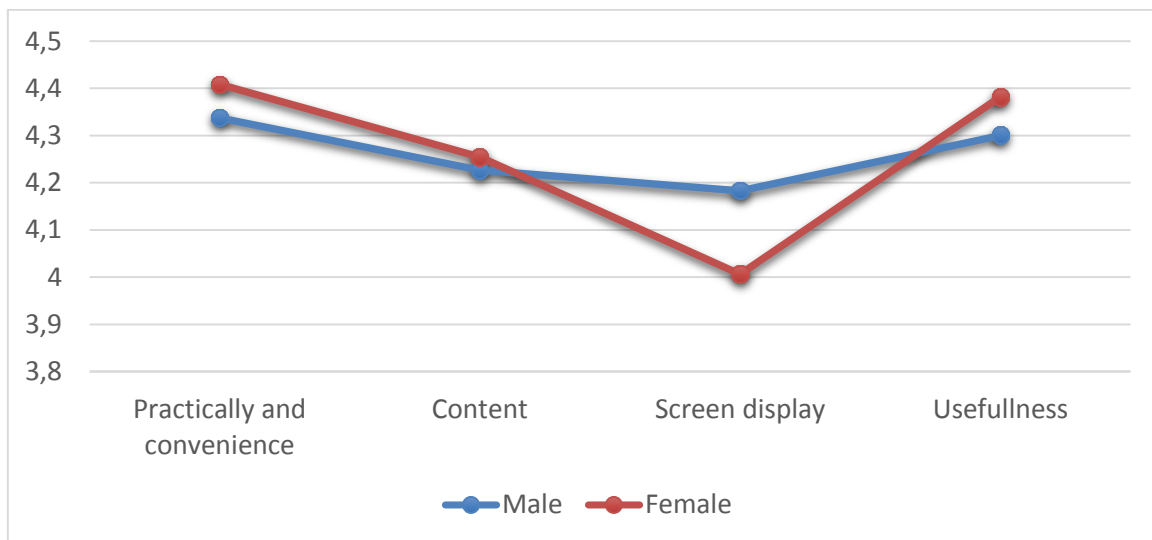

Figure 4. The results of the feasibility test

The results of the feasibility test showed that AnBuso software, developed in this study, was considered very feasible by the respondents. Of the total respondents, 51% said that it is feasible, 46% stated it is very feasible and the rest stated it is quite feasible. This shows that the software has a high usefulness in helping teachers perform item analysis.

The aspect that gets the highest score is its practicality or convenience, and usefulness. The lowest rated aspect is the visual graphic display (see Figure 4). As the respondents reported, the AnBuso is very practical to assist them analysing test items and it really provides them with complete information. The software is able to analyze not only multiple choice test items but also essay tests. The results of the test-item analysis are presented in the forms of report, therefore it is very easy for teachers to interpret them.

IV. CONCLUSION

This study successfully developed AnBuso, which can be used by Indonesian teachers to analyze items practically. This software is developed on Microsoft Excel program which contains data input speadsheet, data processing spreadsheet, analysis report sheet in tables and chart. AnBuso is considered as a very feasible software by the teachers regarding its practicality and convenience, usefulness, and contents, while its screen display is considered only feasible. As AnBuso is very feasible and very useful, it should be disseminated more widely to teachers. This software needs to be upgraded periodically in order to fulfill teachers' future need.

V. REFERENCES

1] J. T. Guliker, T. J. Bastianes and P. A. Kirschner, "A Five-Dimensional Framework for Authentic Assessment," *Educational Technology Research and Development,* vol. 52, no. 3, pp. 67-86, 2004.

2] M. K. Russell and P. W. Airasian, Classroom assessment: concepts and applications (7th edition), New York: McGraw-Hill, 2012.

3] J. E. Raymond, C. S. Homer, R. Smith and J. E. Gray, "Learning through authentic assessment: An evaluation of a new development in the undergraduate midwifery curriculum," *Nurse Education in Practice,* vol. 13, no. 5, pp. 471-476, 2012.

4] T. H. Bers, "The role of institutional assessment in assessing student learning outcomes," *New Directions for Higher Education,* vol. 141, pp. 31-39, 2008.

5] R. J. Stiggins and J. Chappuis, An introduction to student involved assessment for learning. Sixth edition, Boston: Pearson assessment training institute, 2012.

6] Republika, "Implementasi kurikulum 2013 masih dibayangi banyak masalah," 2014. [Online]. Available: www.republika.co.id. [Accessed 18 November 2014].

7] Metronews, "Ini delapan masalah dalam implementasi kurikulum 2013," 2014. [Online]. Available: News.metronews.com. [Accessed 19 October 2014].

8] Tempo, "Problematika implementasi kurikulum 2013," 2013. [Online]. Available: www.tempo.co. [Accessed 10 July 2013].

9] Susilowati, "Kurikulum 2013, 87 persen guru kesulitan cara penilaian," 2013. [Online]. Available: http://unnes.ac.id. [Accessed 18 November 2014].

10] Tempo, "Kurikulum 2013, Apa Saja Kendalanya?," 2014. [Online]. Available: www.tempo.co. [Accessed 16 August 2014].

11] T. W. Banta and C. A. Palomba, Assessment essentials: Planning, implementing, and improving assessment in higher education, San Fransisco: Jossey-Bass, 2014.

12] L. Darling-Hammond, Next generation assessment: Moving beyond the bubble test to support 21st century learning, San Fransisco: Jossey-Bass, 2014.

13] M. D. Miller, R. L. Linn and N. E. Gronlund, Measurement and assessment in teaching (tenth edition), New Jersey: Pearson Education Inc., 2009.

14] Puskur, Model Penilaian Kelas Kurikulum Berbasis Kompetensi, Jakarta: Balitbang Depdiknas, 2008.

15] L. W. Anderson, Classroom assessment: enhancing the quality of teacher decision making, New Jersey: Lawrence Erlbaum Associates, Inc., 2003.

16] T. Kubiszyn and G. D. Borich, Educational testing and measurement: classroom application and practice. 10th edition, New Jersey: John Wiley & Sons, Inc., 2013.

17] J. C. Nunnally and I. H. Bernstein, Psychometric theory (third edition), New York: McGraw-Hill, Inc., 1994.