Improving Teacher's In Developing & Analyzing Made Test Through Follow-Up At CNC Machine Training

Conference Paper · October 2019

3 authors, including:

Eko Indrawan
Situs Resmi Universitas Negeri Padang

Abdul Aziz
Situs Resmi Universitas Negeri Padang

Improving Teacher's In Developing & Analyzing Made Test Through Follow-Up At CNC Machine Training

Yufrizal A, Eko Indrawan, Abdul Aziz
Jurusan Teknik Mesin, Fakultas Teknik, Universitas Negeri Padang
Padang, Indonesia
yufrizal639@gmail.com

***Abstract*--This research aims to improving teacher's in developing &analyzing made test throughfollow-upat cnc machine training. The research method used was qualitative approach through action research. The technique of data collection with the involvement of researchers, among others: discuss, analyze and provide feedback on the question of the objective test the teacher-made (process and product). Data collecting tool that is used to view the processes and products objective test question used sheet checklist Litbang Jakarta 2000/2001 combined with Oriondo and guidelines for analyzing the problem related to the index lurch, different resources, and functioningdistracters. The results of the analysis of the data indicator that at each stage of the activity cycle, the cycle I and cycle II, III, has resulted in an increase against the ability of teachers in making an objective test the teacher made problem, where at each cycle I, obtained numbers < 60%, 80%, Cycle, Cycle III, > 80%. That figure, is information about the number of grains of matter that either meets the criteria set out objective tests of each problem is generated or made of teachers at each cycle, the conclusion that the action research activities with feedback, has been able to increase the ability of teachers in designing CNC machine, create and analyze the problem of objective tests of artificial teachers.**

Keywords: improving teacher's, developing &analyzingmade test, folloe-up

I. INTRODUCTION

One of the teachers in carrying out the duties of her profession is doing an assessment or evaluation of student learning outcomes. Teacher quality needs to be improved in terms of learning and assessment of student learning outcomes, one of which through improved quality of learning and student learning outcomes assessment and improvement of the quality of educators strategic positions and will positively impact against the results of his studies, in addition to other positive impacts.

Teachers conduct the assessment or evaluation of the results of the study, mean teachers do a decision-making process about students, whether students succeed or fail in achieving the learning objectives. Teachers need to make the decision-making process carefully. The wrong decision making will be detrimental to various parties, among others; students, parents, community and school. Therefore, in evaluating or assessing student learning outcomes, the teacher should do it properly so that it will not adversely affect the various parties, and was able to describe the ability of the students.

In doing the assessment of student learning, then one of the teachers is to make test problem. To get a good quality or problem, then the teacher should do good test design. Steps in designing the test begins with determining the purpose test in accordance with the purpose of teaching, setting up a table of specifications, select the format of the grain, the grain, the grain edit writing tests, perform tests, perform analysis, determining the validity of a grain test, determine the reliability test, and interpret test scores. This is the procedure or the General steps in the design and manufacture of test results or the development of learning. Whether teachers have done and are able to design and make the question of the test in accordance with the procedure and the evaluation principles or good?.

Answered the question the study early. Based on preliminary results of a study that was done against some of those teachers who are in the field, particularly in SMK in the field, retrieved a picture that teachers tend to have not made a matter on the basis of the expected process. In General, teachers tend to make no problem, but taking a matter of printed books, book a work sheet (LKS), student and teacher handbook. It describes the teacher has not been paying attention to the purpose test in relation to the learning objectives, aspects that are measured, the validity of the test, reliability test grains, difficulty level (index) a matter of distinction, functioning, distracters carefully. Finally produce tests that are less good. This can be seen when held exams try out held at SMK Negeri 1 Padang is on one of the vocational subjects. Many of the exam questions objective multiple choice shaped given, cannot be judged (given a score), because it has the items and answers wrong. Means the process of designing tests is less good, because the edit problem does not run.

Preliminary study and dialogue. There are problems of teachers identified in the performance assessment of student

learning. Especially in terms of creating or developing the reserved good learning outcomes tests.

Development of an objective test in General through a spate of measures undertaken in a row. Each step is done based on the next step. Development of raw tests are generally carried out through the following steps [14]:

1. Development of test specifications
2. Writing test question
3. Item review
4. Testing the details matter in empires and
5. The final form of the test Administration

1.1 *Development of test specifications*

The nature of the beginning steps in planning the development of a test is a test specification set, i.e. a description which indicates overall quality and characteristics that must be owned by a test that would be developed.

The activity plan of the test specification is essentially a decision-making process. Any decision should be taken based on the considerations regarding various things, such as for example the goal will be achieved, how later tests were used, the influence of a variety of alternative to the validity and reliability tests, and so on. Considerations that generally will be about the following

a. Define the General objectives and the requirements of the test
b. Compose ”blue-print”or test lattice which contains in particular the scope as well as pressure tests and the parts.
c. Select the types of problem
d. Determine the degrees of difficulty problem and its distribution
e. Determine the number reserved for the entire test and for each of its parts.
f. Determine how to compile problem in final form, and
g. Item writing and item review

1.2 *Writing test question*

A common thing that test results of learning (psychological) that consists of a number of questions compiled according to a specific system. Therefore the operational activities in translating the first test specification is writing about. The writing problem is essentially a kind of creation/creations. Writing skills matter more a tips/art than science, although it could not be denied knowledge of the technical stuff in the writing of a matter will improve the ability of writing about. Writing skills matter demands a combination of various special abilities that can only be developed slowly-gradually through practice and experience, by always considering the signs of technical writing about that. Specific skills that must be developed to the extent that adequate. Special abilities of the outline is:

a. mastery of knowledge tested
b. awareness of the values underlying education
c. understanding the characteristics of individuals who are tested
d. the idea language proficiency
e. engineering Mastery writing problem
f. awareness of strengths and weaknesses in writing problem

1.3 *Item Review*

The question has been written with care based on various considerations cannot be casually perceived problems which are already good. Problems still need to be tested quality theoretically. This theoretical question testing is technically called the study a matter or an examination problem (item review). Review of the problem is the evaluation of questions written on the basis of the opinion of the Professional (professional judgment). Three things to be a test in the study about it, namely: (a) the suitability of the content problems with things that will be tested (validity content), (b) the suitability of the problems were with the terms psychometric and (c) the precision and accuracy of the formulation problems such that.

1.4 *Testing the details matter in empires*

The question of who had graduated from testing through a review of the matter in theory are already good. But whether such a test (the test that is in theory good) in the field later would also be nice. To ensure it, needs to be done the test empires testing. In the matter of testing in empires it conducted trials, grain analysis problem (difficulty levels include reserved power and distinguishing items), selection or election matter, namely choosing problems that will be entered into the final form of the test device and which are forced to set aside. Where this problem in the election, according to the classical model of calculation of two parameters, namely, the difficulty level and index discrimination.

1.5The final form of the test Administration

After compiling a matter into the final shape of that test, test development activities has resulted in tests that in theory good and empirically is also good. The next living how tests are used. The teacher has a duty and a fairly complex job. To be able to run errands and her work, then a teacher should have the ability. With regard to the ability of professional teachers. A person's ability can be seen from the results of his work. [8] says that the capabilities, expertise, and results of the work can be assessed using a check list (checklist). More [9] suggested that the results of the work (performance) one can see by the way or tool that can be used as an instrument, namely (1) the list check (checklist), (2) the record anecdotes (anecdotal record), and (3)floor scale (rating scales). A check list is a list to assess the results of one's work in a procedure. Check list contains the description of the proper measures (appropriate steps or correct) and steps that are inappropriate (inappropriate steps or errors), as well as statements about "Yes" or "no". A

checklist of products focuses on the qualities that have made someone (teacher or student). In assessing the ability of teachers to make the question used a checklist of guidelines for composing multiple choice question [3].

Feedback given someone to others, friends of the associate, colleague, and friend to establishing. Next [9] said that feedback can be given to not only the students, but teachers also need to get feedback, i.e., about how he was doing teaching, including ways to evaluate and how their students learn.

The form of the test which can be used by teachers in assessing learning progress is the objective test is multiple choice and essay tests. For the objective test, a good test or quality test is a test that can meet the requirements of (1) validity, (2) reliability, (3) item difficulty index, (4) discriminate index items, and (5)functioning distracters. Regarding the validity of the test, there are three forms of the validity of the test, namely content validity, invalid construct validity, and the validity of the criteria. The validity of the content related to the suitability of the details of the test with measured objectives. Procedures to assess the validity of the content includes: (1) explain the subject and subject matter, (2) set the subject and subject matter as measured by every grain of tests, and (3) match the details of the tests with the subject and the subject being measured. To determine the validity of the content of assessment used panelists [4] and validity of grains tested test using the formula of Point Biserial [1]. Meanwhile the validity of invalid constructs is a type of validity that shows the extent to which the tests reveal trait (nature) or the theoretic invalid constructs measured [13]. To establish the validity of invalid constructs can be done with logical and statistical analysis [15].

A test is said to be the test of reliability in a consistent information about examine [15]. There are several formulas that can be used to determine the coefficients of reliability tests. This depends on the nature of the test score, the dichotomy or continuum. Therefore, the objective test has a score that is the dichotomy, then the appropriate formula used is KR 20 developed by Kuder and Richardson [6].

$$\boldsymbol{r_{11}} = \left(\frac{\boldsymbol{k}}{\boldsymbol{k-1}}\right)\left(\frac{\boldsymbol{V_t - \sum\ pq}}{\boldsymbol{V_t}}\right. \quad \ldots\ldots\ldots\ldots\ldots\ldots (1)$$

After the analysis of validity and reliability tests, then proceed with the item analysis to test the index lurch granules, and different power functioning distracters. Index of difficulty of the grains is the index shows the percentage of test participants who answered correctly and the formula used is:

$$\boldsymbol{p} = \boldsymbol{B}/\boldsymbol{T} \quad \ldots\ldots\ldots\ldots\ldots\ldots (2)$$

where: B = the number of participants and T = total participants test [15]. [11] interpret the index lurch rounds as follows: 0.81 – 1.00 = very easy, 0.21 – 0.80 = moderate, and 0.00 – 0.20 = very difficult. In the meantime [10] citing the opinion of Tuckman who advocated that the index lurch that grain is between $0.33 \leq p \leq 0.67$. Grains with the index $p = \leq 0.33$ is hard grain, $p = > 0.67$ was the grain is easy.Power analysis done subsequent different grains (reserved) tests. Power is the ability of grains the difference with his score can distinguish the respondents-capable capable of high and low. Formula to calculate the different grains of power are:

$$\boldsymbol{D = PH - PL} \ldots\ldots\ldots\ldots\ldots\ldots (3)$$

Where: the PH is the proportion who answered true tests for grain groups up, and PL is the proportion who answered a grain test for lower group [15]. Different grain power index has a range of -1 to +1. A value of +1 indicates perfect difference power, while a value of 0 indicates no power is different on the top and bottom groups. In the meantime the value –1 may occur if the test-taker on the group below can answer correctly a grain of matter, whereas the test-taker on the group over the answer wrong grain tests. [7] said that the power index (ID) = 0.4 to 1 is very good (very good), ID = 0.30 to 0.39 is good (good), the value of ID = 0.20 to 0,29 is small (marginally), ID = 0.00 to 0,19 is revised (revised), and the ID values below zero discard or rewrite.

Then a distracters can be retained if it meets the conditions: (a) answer keys must be selected more by the top of the Group on the Group down, (b) foils should be selected a minimum of 2% of the overall test-taker and selected a minimum of 5% of the Group down, (c) different power index, must be negative [8].

In addition to the above requirements that must be met, in the manufacture of the test multiple choice objective shape, then there are a few things that need to be consideration for the teacher in making grain-shaped test objective relating to the stem (statement) and an alternative answer. [9] suggest for improving the quality of stem from objective tests, there are 4 point to do and 5 point to avoid. For example, the controls do the vocabulary and sentence structure (do not use long sentences, vocabulary, and sentence structure are complex), whereas to avoid, avoid taking the same word with text book. Next [9] suggest for improvement the quality of alternate answers, there are 7 points to do and 7 point anyway to avoid. For example, all the alternatives to do homogeneous, whereas to avoid, avoid the alternatives answers that overlap.

To determine the success rate of teachers in making the problem, in addition to fulfilling the requirements of validity, reliability, the index lurch granules, functioning distracters, a statement of the question, and the answers, then it will also use the guidelines or criteria of multiple choice question according to [3] and [11].

The ability of teachers can be improved through the education, research and training. The existence of this class action research is a means for teachers given information (knowledge), researching and trained about making or crafting a grain test in accordance with the guidelines or a guide. Many of the guidelines that can be used in making

the rounds (item) tests, but guidelines used according to Litbang Jakarta 2000/2001.

Knowledge about creating test items from the aspect of material, construction tests, and language used according to the guidelines. After that teachers make home made test items. Then examined, guided and trained to improve test items are not in accordance with the guidelines. If this is considered eligible, then thetestis givento students. The results of the students' answers were analyzed to determine the validity of the test, the index of difficulty, different power and functioning distracters. Before the teacher to analyze item are given aknowledge of how or the formula used for the analysis.

In the beginning was the teacher will have difficulty in making and analyzing grain test in accordance with the guidelines. A criterion for the desired success has not been high. But after a few rehears alsoraction (3 cycles), then the targetis expected tobe achieved. Success criteriato be achieved adjusted with the principle of mastery learning. Interms of making about the test, the success criteria is greater than 80%. For the validity of the content used expert assessment/rater which is a collaboration of researchers with partisans. Mean while to analyze the validity ofthe test, index of difficulty, different power and functioning distracters used criteria of success above 80%. Success criteria used see Table1.

Table1.Success Criteria Ability Master

No.	CriteriaTheoretical	AspectsAssessed	Assessment Criteria
1.	Constructiont est	Correspondence betweenthe waysteachers makeabout thetestin accordancewith the guidelines(process)	> 80 % (succeed) < 80 % (before)
2.	The validity of the test	Validity of the content(product) Value r-pbi>rcritical(validity) Value r-pbi<r critical(drop)	Expert assessment/ rater Collaboration researchers with participants > 80 % (succeed) < 80 % (succeed)
3.	The difficulty index the question	Soal: Sukar (p = < 0,30) Sedang (p = 0,30 – 0,70) Sulit (p = > 0,70)	If the problem difficulty index(p) =0.30 to 0.70, meaning 80% already achieved success
4.	Different power index	Baik (D > 0,40) Cukup (D = 0,30 – 0,39) Kurang (D = 0,20 – 0,29) Jelek (D = > 0,20)	If the indexis different about D=0.30or>0.30, has achieved a success rate of 80%
5.	Functioning distraktor	At least 2% of Respondents answer each option	> 80 % succeed < 80 % (before)
6.	the ability of teachers Creating and analyzing the question	In accordance with guide lines or guides	> 80 % (succeed) < 80 % (before)

Outcomes of research objectives are to: 1) Determine the ability of teachers CNC in designing and creating multiple choice test questions in terms of process or procedure, 2) Knowing CNC teachers' ability to analyze objectively shaped test item (item analysis) artificial covering; validity, reliability, power index difference, level of difficulty and functioning **distraktor, 3)** improve the ability of teachers CNC in designing, making and analyzing the shape of an objective test questions in terms of the product through the feedback, in order to obtain quality problems.

II. METHOD

This research is action through a qualitative approach, which seeks to multiply the problems based on real situations that occur in the field. Therefore, the action and the data obtained is maintained naturally as possible without any element of intervention on the part of researchers. Natural data then reflected through agreements with the participants of the study.

Where the research SMK Negeri 1 Padang. Subjects were teachers, especially teachers of vocational subjects. These subjects are tested through the national final examination (UAN). Therefore, there are constraints of time, cost, and energy research, the research focused on the subject of teacher CNC course, with the planning, preparation, execution, seminars and preparation of final report.

This research was conducted in three cycles and each cycle consists of four activities, namely: (a) planning, (b) measures, (c) observation and (d) reflection. Procedures action research activities on the first and subsequent cycles can be seen in Table 2 below:

Table 2. Action Research Procedure Research Activities on Third Cycle

No	Kegiatan	Cycle I	Cycle II	Cycle III
1.	Planning	a. Plan learning materials b. Choosing a method c. Creating tools (media) d. Designing a data collection tool e. Designing tasks f. Setting up the interview guide g. Setting up the observation sheet h. Prepare a reflection sheet i. Setting up a data processing device j. Setting up a feedback sheet	Recycling for the implementation of the first cycle of action, focusing on repair	Recycling for implementation action cycle second, focus on repair
2.	Action	a. Provide theoretical information about evaluation and measurement in general, and especially make a good question b. A meeting with partners and invites teachers make improvements in the field of evaluation c. Receiving various inputs from partners teacher d. Discussing on how to improve the quality problem e. Preparation of matter f. The collection of the data g. Data processing	sda	sda
3.	Observation	a. Observing teachers design and create questions b. Observing comments or responses that lasts c. Noted weaknesses conducted by partner teachers d. Observing the role of peers e. Observing the role of principals	sda	sda
4.	Reflection	a. Discussing about the achievement, obstacles and drawbacks encountered b. Discuss possible contributing factor c. Draw/set conclusions d. Discuss solutions as a corrective action (how, where, and when) in the next cycle.	sda	sda

During the implementation phase information, the data obtained ist he feedback for further action. There are several possibilities of change. First, in accordance with the action plan and has not shown any change. The consequence is that implementation should proceed according to a predetermined plan from the beginning. Second, the lack of action in accordance with the plan and has not shown any change. Consequently, the action plan be revised or corrected. Third, the action is not according to plan but does not show the changes that lead to the achievement of goals. Consequently, planning and implementation of measures need to be revisited. Fourth, thelackof actionin accordance withthe plan, but showedno changesthat leadto theachievement of objectives.Consequently, actionsneed to be refined. For more details,this can be seenin Table3 below:

Table3. Framework Model Compliance Action, Change, Actionsand Consequencesin research

No.	**SuitabilityActions**	**Change**	**Consequence**	**Description**
1.	Less appropriate	Belum	repair	next cycle
2.	Less appropriate	Ada	repair	Sda
3.	Corresponding	Belum	repair	Sda
4.	Corresponding	Ada	no repair	stopped cycle

Based on Table 3 above it can be concluded that the fit between the action plan with the expected changes at all to determine whether or not the cycle continued.Data collection techniques is by researchers involved in the activities undertaken. Involvement of researchers includes activities such as: discussing the draft about, write about that in accordance with the purpose of teaching and aspects to be tested, select the language (vocabulary and structure) about the consideration of the advice of experts in writing of items and alternative answers, go to class at the time of the exam held, analyze and process the data results of the exam, and set decisions about student achievement.

Data collection tool used to look at the processes and products used to make about the check list sheet according to [3] combined with Oriondo and guidelines to analyze problems related to difficulty index, different power and functioning distraktor.

III. RESULTS AND DISCUSSION

The research was conducted or carried out by the teacher (teacher partners) who taught in Class XI SMK 1 Padang, on subjects CNC first semester of the school year 2016/2017. In the implementation of actions or activities are divided into several cycles. The results of research or action in each cycle can be described as follows:

Table 4. Results of Analysis of Problem Test Objective Artificial Master Cycle I

No	Rated Aspect	Theoretical Criteria	Porsentase Achievement
1	Construction test	Conformity between ways teachers making test questionin accord ancewiththe guidelines	70%
2	Test Validity and Reliability Tests	Validity Tests Nilai r-pbi > r kritis (valid) Nilai r-pbi < r kritis (drop)	r-pbi = 0,263
3	The difficulty index of question	Soal Sedang (p = 0,30 – 0,70)	60%
4	Power Difference Index	Cukup (D = 0,30 – 0,39)	70%
5	Functioning distractor	At least 2% of Respondents answer each option	60%
6	Ability Master Creating and analyzing question	In accordance with guide lines or guides	60%

Based on the information or data obtained by further discussions between teams of researchers and teachers as partners on the achievement of the feedback to take further action. From the data obtained as shown in Table 1 it can be concluded, the teacher's ability to make objective tests are still
weak, lacking action in accordance with the plan and the action has not showed changes that lead to the achievement of goals. As its consequences and the implementation of the action plan needs to be reviewed and reversible, and the activities continued into the second cycle

Table 5. Results Analysis of Problem Test Objective Artificial Master Cycle II

No	Rated Aspect	Theoretical Criteria	Porsentase Achievement
1	Constructiontest	Conformity between ways teachers making test question in accordance with the guidelines	80%
2	TestValidity and Reliability Tests	Validity Tests Nilai r-pbi > r kritis (valid) Nilai r-pbi < r kritis (drop)	r-pbi = 0, 4
3	The difficulty index of question	Soal Sedang (p = 0,30 – 0,70)	85%
4	Power Difference Index	Cukup (D = 0,30 – 0,39)	80%
5	Functioning distraktor	At least 2% of Respondents answer each option	75%
6	Ability Master Creating and	In accordance with guide lines or guides	80%

	analyzing question		

In the third cycle, planning activities and actions carried out the same as in the planning stage and the action taken in cycles I and II. Likewise, observation and reflection activities under taken in the second cycle. The findings of each phase of activity in the third cycle, as reflected in Table 6 below. From information Table 6 it can be concluded that the action was in accordance with the plan and show the changes that lead to achieving the goal, and the cycle can be stopped.

Table 6. Results of Analysis of Problem Test Objective Artificial Master Cycle III

No	Rated Aspect	Theoretical Criteria	Porsentase Achievement
1	Construction test	Conformity between ways teachers making test question in accord ancewiththe guidelines	85%
2	Test Validity and Reliability Tests	Validity Tests Nilai r-pbi > r kritis (valid) Nilai r-pbi < r kritis (drop)	r-pbi = 0, 4
3	The difficulty index of question	Soal Sedang (p = 0,30 – 0,70)	85%
4	Power Difference Index	Cukup (D = 0,30 – 0,39)	80%
5	Functioning distractor	At least 2% of Respondents answer eachoption	80%
6	Ability Master Creating and analyzing question	In accordance with guide line sor guides	85%

Research conducted on the ability of teachers to create or write Problem Test Objective Artificial Master himself. Conducted in three cycles with multiple stages, produce which can be explained as follows:

In the first cycle; obtained information about the ability of teachers in compiling and analyzing objective tests, generally still below the specified criteria (<70%), both in processes, procedures and products. This indicates that the ability of teachers to create questions in the form of an objective test is still weak (less good). According to researchers weakness was caused primarily; Teachers lack training and have yet to understand and absorb a lot of information that has been provided, include: knowledge of the evaluation, including ways to make an objective test that is good and true, and the knowledge of how to analyze it. However, after discussion, describe the information that has been given, and then provide feedback (feedback) of the errors were found and given a proper example, and activities continued in the second cycle. Turns out affecting changes increase the ability of teachers to create an objective test questions. This is evident from the achievements gained increased to 80%, meets the set criteria, but have not achieved> 80%.

The increase that occurred in the second cycle, according to the researchers was in addition, has increased knowledge and understanding and skills of teachers in the evaluation, especially making an objective test questions, through feedback and the discussions held at the beginning of the second cycle. Also the teachers because they already have the knowledge and skills toconduct an analysis of different power levels, level of difficulty, significance distraktor, the validity and reliability of the matter objectively made and take advantage of the analysis conducted. After the feedback and discussion at the end of the second cycle. To believe in the mastery of knowledge/understanding and skills gained from discussion and feedback provided the research team is correct/right, then the activity again (the third cycle). Apparently in this third cycle, the results occur again increase the ability of teachers to create/write objective tests. More than 80% of objective test questions are prepared to comply with the guidelines established criteria. The success of the teacher is inseparable from the desire (motivation) of teachers concerned to improve their competence in learning activities. In this case, especially in making an objective test questions. Their contention, made a good question following the evaluation guidelines is difficult and not very useful for improving the quality of learning is to be imprecise.

IV. CONCLUSION

From the study of the teacher's ability to create or to develop test questions in the form of objective, can be summed up as follows: (1) The ability of teachers CNC in making shaped test questions objectively, in terms of processes and or procedures based on the Guidelines for Preparation of Multiple Choice Questions, prior to this study, is very low. Only 60% of the teachers about the artificial objective which meet the criteria Guidelines for Preparation of Multiple Choice Questions. For most teachers at SMK 1 does not obtain information about the guidelines for the preparation of test used. Schools no set guidelines used. (2) Ability to analyze problems Guru CNC-shaped test objective (products), before the research performed is still very weak. Most teachers have never and do not know how to conduct an analysis of objective test questions were made, particularly in analyzing the validity and reliability of a test. (3) Efforts feedback about objective tests carried out on teacher-made both in terms of processes, procedures, and products, has been able to improve the ability of teachers to create and analyze a matter of objective tests of artificial teachers, is reflected in the results seen from each stage of the cycle is done of>70% in the first cycle, 80% in the second cycle, and>80% in the third cycle.

Based on the finding so fresearch that has been conducted onthe ability of teachers to create questions objective tests suggested as follows: (1) Teachers, making objective tests, shall pursue the procedures and ways that really make upa matter of objective tests, according to the evaluation of science and or guidelines established. (2) The teacher letaugment and improve the science evaluation, particularly in analyzing the matter, through education and training, discussions among peer sand others. So asa matter ofobjective testhas been made through the analysis, so itis expected to measure students' actualability. (3) Master let discuss anymatter of objective tests were made, withpeers,so thatthe feedback that occurs from the discussion, it can improve thequality of the tests and the ability of teachers them selves.

References

[1] Algina, James and Linda Crocker. 1996. *Introduction to Classical and Modern Test Theory*. Florida: Holt Rinehart and Winston, Inc.

[2] Departemen P&K. 2008. *Penelitian Tindakan* (*Action Research*), Bahan Penelitian. Jakarta: Proyek Peningkatan SMU.

[3] Depdiknas. 2000/2001. *Bahan Penataran Pengujian Pendidikan.* Jakarta: Pusat Pengujian-Pusjian.

[4] Ferguson, George. A. 1992. Statistical Analysis in Psychology and Education. Tokyo: McGraw-Hill.

[5] Gronlund, Norman. E. 2002. *Measurement and Evaluation in Teaching*. New York: MacMilan Publishing Company.

[6] Guilford, J. P. 1982. *Psychometric Method*, New York: McGraw-Hill Company, Inc.

[7] Hopkins KD, Stanley JC and Hopkins R.2005. *Educational and Psychological Measurement and Evaluation*. Boston: Allyn and Bacon.

[8] Linn, Robert L. 2001. *Educational Measurement*. New York: MacMillan Publishing Company.

[9] Nitko, Anthony. J. 1996. *Educational Assessment of Students*. New Yersey: Prentice Hall Inc.

[10] Naga, Dali S. 2002 *Pengantar Teori Sekor Pada Pengukuran Pendidikan*. Jakarta: Gunadharma.

[11] Oriondo, Leonara Loyola and Eleanor M. Dallo Antonio. 1998. *Evaluating Educational Outcomes*. Manila: Book Store.

[12] Phopam, James W. 1995 *Classroom Assesment*. *What Teacher Need to Know*. Los Angeles: Allyn and Bacon.

[13] Wiersma, William and G Jurs. 1990. *Educational Measurement and Testing* Boston: Allyn and Bacon