CHAPTER V

SUMMARY, CONCLUSION AND RECOMMENDATIONS

5.1.SUMMARY AND CONCLUSION

Nowadays the ability to make a good test becomes an urgent requirement for a teacher. It has been taken for granted that our country now demands a lot of truly qualified people to catch up with the development of science and technology. To determine whether a testee is really qualified or not in his field can be found out from the results of his tests. ; therefore, the role of testing in education world becomes more and more important than before. One of the criteria to be a good test is that it should have predictive validity - the results of the test should really reflect the true ability of the testees so that they can be used to judge their future achievements.

However the writer still finds that many results of tests do not reflect the true ability of the testees. This view attracts the writer to conduct an observation study to the 1989 Structure I final semester test. In this study, the writer is curious to find out whether the test has the degree of predictive validity or not.

In order to find out the answer, the writer took two kinds of data from the student academic records that were the scores of 1989 Structure I test and the scores of

1990 Structure II test. After obtaining the data, the writer examined the correlation between the two tests' scores, calculated the deviation of the estimated scores from the actual scores and examined the significance of the regression coefficient in making prediction.

The results of the calculation show that :

1. The correlation coefficient of the 1989 Structure I test and 1990 Structure II test is 0.7758 . The coefficient is bigger than r table and it means that the two tests are significantly correlated.
2. Thirty-two students got higher scores than the writer estimation and forty students got lower scores than the writer's estimation. In this calculation, the writer also found that some students' scores (sample number 30, 36, 38, 44, 45, 46, 55, 64, 68, 70) deviate much higher than the estimation (more than nine) and some students' scores (sample number 3, 13, 19, 20, 35, 42, 47, 62, 66, 69, 72) deviate much lower than the estimation (less than nine) - the figure nine is taken from the standard deviation of the estimated scores. The great deviation means that the students' scores on the Structure I test do not reflect their true ability at all.
3. Regression coefficient calculated is 1.017037 and the ratio for the regression coefficient is 10.2854 (it is

bigger than t table). It means that the regression coefficient is significant in playing a role in making prediction of the students' achievement in the 1990 Structure II test.

Although there are 29.2 % students' scores deviating too great from the writer's estimation, the writer can draw a final conclusion that the 1989 Structure I test has a fairly high degree of predictive validity and it means that it has fulfilled one of the criteria to be a good test.

5.2. RECOMMENDATIONS

Based on these findings, it is recommended that a teacher should be careful in measuring his students' achievement. To know whether a test has the degree of predictive validity or not, a teacher looks only at the results of the test. Since there are many factors which can influence a teacher to make a measurement error, he should be able to identify, to prevent and to estimate the errors which can happen.

First of all, a teacher should see the content of a test and a table of specifications is needed in this case. If the information needed is about the student's understanding of structure concepts, the questions in the test should be structure questions. Furthermore, if a

representative sample of different topics is not included in the test questions, then the information gained will be questionable.

Second, a teacher should check the item difficulty of the test. If the test items are too easy or too hard, they will likely yield unreliable information since it can not differentiate between slow students and good ones. If an item is very easy, all of the students are supposed to be able to answer the item correctly. On the contrary, if an item is too difficult, most students will probably score low and it means that the item also can not be used to distinguish various level of ability.

Third, a teacher should give a clear direction to his students so that they can be certain how they are expected to respond to the items. If a student is unsure of the instructions in the test sheet and uncertain what a question is asking , he will be confused and it will also influence his score.

Fourth, the time of the test administering should also be considered. A test must, for example, be long enough so that a large sample of information needed can be obtained and so that the results of the test can be stable over time.

Fifth, the conditions under which a test is given must be taken care of. For example, if a test is given in

a room where there are numerous distractions (e.g. people walking through the room or lawn mowers running outside the window), the accuracy of the scores may be affected.

Finally, a teacher should know about each student being evaluated for example about his general health. Although the student's general health is not always easily determined by the teacher, it is an important factor in the accuracy of any evaluation which the teacher makes. If a student seems ill, a teacher should ask the student what trouble he feels is. If it is too hard for him to continue doing the test, the teacher should give another test with the same difficulty on the other day - if the illness from which the student suffers is not too serious, it is expected to give the other test not too long from the time of the preceding test.

From the writer's experience, there are usually rotating teachers teaching a certain subject in the English Department of Widya Mandala University. It means that a teacher who teaches a subject cannot be certain that he will teach the same subject again at the new semester. Since validity of a test is specific to a particular use, a valid test will not be valid anymore if it is used by different teachers. Furthermore, a specific valid test cannot be used as a model to be administered in another occasion. Before ending this thesis, here the

writer wants to remind the readers once again that she made this study with a purpose to see the validity of the 1989 Structure I test as a measuring instrument and not to suggest that structure teachers of English Department of Widya Mandala University should make this test as a model of structure test.

BIBLIOGRAPHY

1. Ausubel, David P., Educational Psychology : A Cognitive View, New York : Holt, Rinehart and Winston.

2. Ausubel, David P. and Robinson F.G., School Learning : An Introduction to Educational Psychology, New York : Holt, Rinehart and Winston, 1969.

3. Chastain, Kenneth, Developing Second Language Skills : Theory to Practice, London : Houghton Mifflin Company, 1976.

4. Cronbach, Lee J., Essentials of Psychological Testing, New York : Harper and Row Publishers, 1817.

5. Corder, S. Pit, Introducing Applied Linguistics, England : Penguin Education, 1973.

6. Draper, Norman and Harry Smith, Applied Regression Analysis, Second Edition, New York : John Wiley & Sons, Inc., 1981.

7. French, The Structure of English, London : The English Language Book Society and Oxford University Press, 1963.

8. Finnochiaro, Mary Ph. D., English as a Second Language : From Theory to Practice, New York : Regents Publishing Company, 1946.

9. Gronlund, Norman E., Measurement and Evaluation in Teaching, London : Collier Macmillan Publishers, 1981.

10. ----------, Constructing Achievement Test, London : Prentice-Hall, Inc. .

11. Guth, Hans P., Words and Ideas, California : Wordsworth Publishing Company, Inc. .

12. Harris, David P., Testing English as a Second Language, New York : McGraw-Hill Book Company, 1969.

13. Heaton, J.B., Writing English Language Tests : A Practical Guide for Teachers of English as a Second or Foreign Language, London : Longman Group Limited, 1975.

14. ----------. Writing English Language Tests, Singapore : Longman Group Limited, 1979.

15. Hopkins, Charles D. and Richard L. Antes, Classroom Measurement and Evaluation, Illionis : F.E. Peacock Publishers, Inc., 1978.

16. Hutchinson, Tom and Alan Waters, English for Specific Purposes : A Learning-Centred Approach, New York : Cambridge University Press, 1986.

17. Johnson, David P., Educational Psychology, London : Prentice-Hall International, Inc., 1979.

18. Lado, Robert Ph.D., Language Testing : The Construction and Use of Foreign Language Tests, London : Longman, Green and Co. Ltd., 1967.

19. Micheels, William J. and M. Ray Karnes, Measuring Educational Achievement, New York : McGraw-Hill Book Company, Inc., 1950.

20. Nunnally, Jum C., Psychometric Theory, New York : McGraw-Hill Book Company, 1978.

21. Parsons, Robert, Statistical Analysis : A Decision Making Approach, New York : Harper & Row Publishers, Inc., 1978.

22. Robinett, Betty Wallace, Teaching English to Speakers of Other Languages : Substance and Technique, Minneapolis : University of Minnesota Press, 1978.

23. Tenbrink, Terry D., Evaluation - A Practical Guide for Teachers, United States of America : McGraw-Hill, Inc., 1974.

24. Walker, Helen M. and Joseph Lev, Elementary Statistical Methods, New York : Holt, Rinehart & Winston Inc., 1969.

25. Weiss / Hassett, Introductory Statistics, Philippines : Addison-Wesley Publishing Company, 1982.

26. Zar, Jerrold H., Biostatistical Analysis, New York : Prentice-Hall, Inc., 1974.