CHAPTER I

INTRODUCTION

This chapter introduces the background of the study, the purpose of the study, a statement of the problem, the scope of the study, the significance of the study, a brief explanation of research method, clarification of terms, and the organization of the paper.

1.1   Background of the Research

Testing is really important in education because testing is used to measure the successfulness of teaching learning process. A test that is conducted at the end of teaching and learning process is called summative test. As Brown (2010, p.7) says summative test is summation of the students' ability/ knowledge after having been given the material in the learning process.

One test that is categorized as a summative test is an achievement test. Achievement test is usually conducted at the end of the learning process to measure the successfulness of the program (Hughes, 2003, p. 13). Achievement test is usually limited to particular material objective in the curriculum. One important achievement test in Indonesia is a final examination that is usually conducted at the end of one school year learning process.

EFL teachers in Indonesia usually develop test instrument for final examination by themselves since there is a regulation in The Ministerial Regulation Number 16 of the year 2007 on academic qualification standards and teacher competencies that says teachers must be able to conduct an assessment and evaluation processes and learning outcomes. It means that teachers must develop assessment instruments and evaluation processes and learning outcome. Thus, it can be said that EFL teachers in Indonesia must be able to develop or write good assessment instrument including the final examination items.

However, writing test items is not as easy as it is seen. It is known that not every good teacher can be a good test maker (Karimi&Mehrdad, 2012, p. 1). It will cause a problem in Indonesia since Indonesian teachers have to develop not only the classroom test but also the final examination. If teachers do not develop a good test, the test which is so important may be invalid.

The common problem that occurs in the test is validity problem. The term validity has a lot of definition and developed over the years. Based on table of validity's development in Hilden (2008), the first terms related to validity are reliability, criterion validity and content validity proposed by Cornbach et al. in 1985, and then it is continuously developing through Messick era, Bachman era, Palmer era, Weir era, until Kane era. The terms proposed are such as reliability, content validity, criterion validity, etc. The most interesting concept is Weir's validity since this concept provides validity analysis before the test event (a priori validity) and validity after test (posteriori validity).

There have been many studies analyzing validity of items. However the literature that discusses validity analysis by using Weir's approach: a priori and posteriori validity is still a few. In fact a priori validity is needed because if one test does not pass a priori validity analysis, it will not also pass posteriori validity analysis.

The latest study about a priori validity is Seyyedrezaei and Birjandi's research that analyzes cognitive validity of general writing task two of IELTS in 2015. Their research is qualitative and quantitative research that is used questionnaire and think-aloud-protocol as its method. The findings show that the test does engages most of cognitive processes that are deemed essential in the target language use domain.

The Seyyedrezaei and Birjandi's research shows that IELTS, one big English testing institution, comprehends of validity theory and can apply it in their test. Then one question arises: if the test is school achievement test developed by teachers, are the teacher able to show a priori validity evidence in their items? Regarding things stated above, the present research is intended to find out the

FADILA ISTIQA SEPTIANA, **2016**
A PRIORI VALIDITY OF SUMMATIVE TEST ITEMS DEVELOPED BY EFL TEACHERS: A Case Study at a Junior High School in Cimahi.
Universitas Pendidikan Indonesia | repository.upi.edu | perpustakaan.upi.edu

result of analysis of a priori validity of test items developed by EFL teachers in Indonesia.

1.2   Research Question

This study is carried out to answer the research question as follows: How is the a priori validity of summative test items developed by EFL teachers in Indonesia?

1.3   Aim of the Research

Referring to the introduction above, this study aims to find out a priori validity of summative test items developed by EFL teachers in Indonesia.

1.4   Scope of the Research

This study is focused on a priori validity of summative test items developed by EFL teachers in Indonesia. By using a priori validity framework proposed by Weir (2005), context validity evidence and theory-based validity evidence is analyzed.

1.5   Significance of the Research

The significance of this study can be seen from three perspectives. Theoretically, this study is expected to enrich the research about assessment topic in EFL countries. This study is also expected to give contribution in giving information about language testing in EFL countries.

Practically, this study is expected to give information to EFL teachers in developing test items instrument in terms of the content validity and the construct validity so that EFL teachers in Indonesia can develop a valid test for class assessment and final examination.

Professionally, the research can help the other researchers who are interested in validity topic by becoming a guideline to support their research.

FADILA ISTIQA SEPTIANA, 2016
A PRIORI VALIDITY OF SUMMATIVE TEST ITEMS DEVELOPED BY EFL TEACHERS: A Case Study at a Junior High School in Cimahi.
Universitas Pendidikan Indonesia | repository.upi.edu | perpustakaan.upi.edu

1.6 Clarification of Terms

The following terms are clarified in order to avoid the ambiguity, misunderstanding and misconception of the case presented in this study.

1. *Summative assessment*

Summative assessment is conducted at the end of learning process such as the end of the term, semester, or year (Hughes, 2003, p. 5) to check students' achievement related to the learning process. A kind of test that is categorized into summative assessmentis achievement test. Hughes (2003, p. 13) states that there are two types of achievement test; final achievement test and progress achievement test. Final examination in this research is categorized as a final achievement test because final examination in this research is administered at the end of the course and the content should be based on the material on the course syllabus. In addition, since Indonesian teachers usually develop their own test for the final examination, the test makers in this research is a teacher who teaches the material and the test takers are the students.

2. *A Priori Validity*

Weir (2005, p.17) states that a priori validity is validity that is established before the test event. Therefore the score of the test is not used in establishing a priori validity evidence in this research since it is not yet collected. A priori validity evidence can be established by analyzing context validity and theory based validity (Weir, 2005, p. 17).

3. *Context Validity*

Term context validity is proposed to refer to content validity since Weir thinks it is a better term for social dimension of language use (2005, p. 19). Context validity relates to how far the items represent all materials in the course. This research analyzes validates context validity by analyzing three aspects suggested by Weir (2005, p. 57-84) such as task settings, task demands, and setting and test administration.

4. *Theory-Based Validity*

Theory based validity is related to the extent how far the test represents the theory underlying the skill that is measured (Weir, 2005, p. 18). In this research, executive resources and executive processing of language skills are identified to know theory-based validity of the test. An interview is also conducted to enrich the data.

5.   Test Items

Test items are the questions in a test that should be answered by the test takers. Format response of items in this research is multiple choice items.

1.7   Organization of the Paper

This paper is organized in five chapters. Each chapter has some subtopics which help the reader to clarify and give details on the information. The content of this research is organized as follows:

Chapter I. Introduction

This chapter encompasses the introduction of the research which covers the background of the research, a statement of the problem, the purpose of the research, a significance of the research, a limitation of the research, research methodology, a clarification of key terms, and organization of the papers.

Chapter II. Literature Review

This chapter presents relevant theories to this research. In detail, this chapter explores some theoretical foundations of test include assessment and evaluation, common test techniques specifically multiple choice test instrument, and validity, specifically Weir's a priori validity including context validity and theory-based validity.

Chapter III. Research Methodology

This chapter presents a discussion on the methodology employed in this research: the description of research design, research subject, research instrument, research procedures, and data analysis.

Chapter IV. Findings and Discussions

This chapter presents the results of data analysis and the discussions. The findings and interpretation of data are provided in one unit by separated the findings of context validity and theory based validity.

Chapter V. Conclusions and Suggestions

This chapter presents the conclusions and suggestions. It includes the explanation of the validity of multiple choice test instruments and some suggestions for the further research in this field.

1.8 Concluding Remarks

This chapter has discussed the background of the study, research questions, purposes of the study, the significance of the study, clarification of terms, and organization of the paper. The next chapter will elaborate the theoretical background as well as related research to be used as the foundation to develop a clear framework of the formulation of the problem to be investigated.