CONSTRUCTING TEST ITEMS : PRE-SERVICE TEACHERS' CHALLENGES AND STRATEGIES

THESIS

Submitted in partial fulfilment of the requirement for the degree of Sarjana Pendidikan (S.Pd) in Teaching English

By

Wilda Shofiarita

D05216034

ENGLISH LANGUAGE EDUCATION DEPARTMENT

FACULTY OF TARBIYAH AND TEACHERS TRAINING

UNIVERSITAS ISLAM NEGERI SUNAN AMPEL

SURABAYA

2020

PERNYATAAN KEASLIAN TULISAN

Saya yang bertanda tangan dibawah ini:

Nama	: Wilda Shofiarita
NIM	: D05216034
Jurusan/Program Studi	: Pendidikan Bahasa/Pendidikan Bahasa Inggris
Fakultas	: Tarbiyah dan Keguruan
Judul Skripsi	: **"Constructing Test Items: Pre-Service Teachers' Challenges and Strategies"**

Menyatakan dengan sebenarnya bahwa skripsi yang saya buat ini benar benar merupakan hasil karya saya sendiri, bukan merupakan pengambil alihan tulisan atau pikiran orang lain yang saya akui sebagai hasil tulisan saya sendiri. Segala materi yang di ambil dari orang lain, hanya sebagai acuan untuk mengikuti tata cara serta etika penulisan karya ilmiyah yang telah ditetapkan oleh jurusan.

Surabaya, 01 September 2020

Yang membuat penyataan

METERAI TEMPEL 2312FAHF317375423 6000 ENAM RIBU RUPIAH

Wilda Shofiarita

NIM : D05216034

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Dipindai dengan CamScanner

ADVISOR APPROVAL SHEET

This thesis by Wilda Shofiarita entitled "Constructing Test Items: Pre-Service Teachers' Challenges and Strategies" has been examined by the advisors for the further approval by the board of examiners.

Surabaya, 25 June 2020

Advisor I,

Fitriah, Ph. D

NIP. 197610042009122001

Advisor II,

Rakhmawati, M.Pd

NIP. 197803172009122002

EXAMINER APPROVAL SHEET

This thesis by Wilda Shofiarita entitled "Constructing Test Items: Pre-Service Teachers' Challenges and Strategies" has been examined by the advisors for the further approval by the board of examiners.

Dean,

Prof. H. Ali Mas'ud, M.Ag. M.Pd.I
NIP. 196301231993031002

The Board of Examiner
Examiner I,

Hilda Izzati Madjid, MA
NIP. 198602102011012012

Examiner II,

Rizka Safriyani, M.Pd
NIP. 198409142009122005

ExaminerIII,

Fitriah,Ph.D
NIP. 197610042009122001

Examiner IV,

Rakhmawati, M.Pd
NIP. 197803172009122002

KEMENTERIAN AGAMA
UNIVERSITAS ISLAM NEGERI SUNAN AMPEL SURABAYA
PERPUSTAKAAN
Jl. Jend. A. Yani 117 Surabaya 60237 Telp. 031-8431972 Fax.031-8413300
E-Mail: perpus@uinsby.ac.id

LEMBAR PERNYATAAN PERSETUJUAN PUBLIKASI
KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS

Sebagai sivitas akademika UIN Sunan Ampel Surabaya, yang bertanda tangan di bawah ini, saya:

Nama : Wilda Shofiarita
NIM : D05216034
Fakultas/Jurusan : Tarbiyah dan Keguruan/Pendidikan Bahasa Inggris
E-mail address : shofiaritawilda@gmail.com

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Perpustakaan UIN Sunan Ampel Surabaya, Hak Bebas Royalti Non-Eksklusif atas karya ilmiah :
Sekripsi ☐ ☐ Tesis ☐ Desertasi ☐ Lain-lain (……)
yang berjudul :

CONSTRUCTING TEST ITEMS : PRE-SERVICE TEACHERS' CHALLENGES AND STRATEGIES

beserta perangkat yang diperlukan (bila ada). Dengan Hak Bebas Royalti Non-Ekslusif ini Perpustakaan UIN Sunan Ampel Surabaya berhak menyimpan, mengalih-media/format-kan, mengelolanya dalam bentuk pangkalan data (database), mendistribusikannya, dan menampilkan/mempublikasikannya di Internet atau media lain secara ***fulltext*** untuk kepentingan akademis tanpa perlu meminta ijin dari saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan atau penerbit yang bersangkutan.

Saya bersedia untuk menanggung secara pribadi, tanpa melibatkan pihak Perpustakaan UIN Sunan Ampel Surabaya, segala bentuk tuntutan hukum yang timbul atas pelanggaran Hak Cipta dalam karya ilmiah saya ini.

Demikian pernyataan ini yang saya buat dengan sebenarnya.

Surabaya, 5 September 2020

Penulis

(Wilda Shofiarita)

ABSTRACT

Shofiarita, Wilda. (2020). *Constructing Test Items : Pre-Service Teachers' Challenges and Strategies.*A thesis. English Language Education Department, Faculty of Tarbiyah and Teacher Training, UIN Sunan Ampel Surabaya, Surabaya. Advisor I: Fitriah, Ph.D. Advisor II: Rakhmawati, M.Pd.

Keywords : Assessment, Test Items, Pre-Service Teachers' Challenges, Pre-Service Teachers' Strategies

Assessment is one of important parts in teaching and learning. It is a process of identifying students' progress, planning future materials, designing interesting learning activities and encouraging students in learning. However, to obtain comprehensive information of students' real ability and progress, teachers should be able to construct test items. It is a technique to measure students' ability, knowledge and performance. This research aimed to explore pre-service teachers' challenges and strategies in constructing test items. There are five pre-service teachers participated in this research. This research involved pre-service teachers who had taken Language Assessment Course, had made test for senior high school, and had score for this unit. To answer the research questions, this study interviewed five participants about some questions such as experience in constructing test items, challenges and strategies in constructing test items. The findings showed that there were five challenges identified in the study. The challenges in constructing multiple choice items are writing good sentences, selecting the options for multiple choice items, balancing the skills within one test, and finding references. Moreover, the challenge in constructing essay item is designing rubric for open-ended questions. One challenges faced in designing rubric is specifying criteria and standards clearly. Writing good sentences refers to pre-service teachers' grammar while writing the test. Then, there are some considerations in writing the options for multiple choice test such as the stem should be meaningful, the stem should be relevant with the material, and etc.From the mentioned challenges, designing rubric for open-ended questions was the most challenge skill that these pre-service teachers encountered.The study also found that ensuring the questions in the test in line with the syllabus and the given materials, writing test online, designing the test clearly and subjective, and writing a clear rubric were pre-service teachers' strategies in constructing test items. This helped them ensure the suitability of test with the students' level and achieve test purposes.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

ABSTRAK

Shofiarita, Wilda. (2020). *Constructing Test Items: Pre-Service Teachers' Challenges and Strategies.*Skripsi. Pendidikan Bahasa Inggris. Fakultas Tarbiyah dan Keguruan. Universitas Islam Negeri Sunan Ampel Surabaya. Advisor I: Fitriah, Ph.D. Advisor II: Rakhmawati, M.Pd

Penilaian merupakan salah satu bagian dalam pengajaran. Ini merupakan proses untuk mengidentifikasi kemajuan siswa, merencanakan materi selanjutnya, mendesain aktivitas pembelajaran yang menarik, dan meyakinkan siswa dalam pembelajaran. Bagaimanapun juga, untuk memperoleh informasi yang luas mengenai perkembangan dan kemampuan siswa, guru seharusnya bisa membuat tes.Ini adalah sebuah teknik untuk mengukur kemampuan, ilmu, dan perbuatan siswa.Penelitian ini bertujuan untuk meneliti tantangan dan strategi calon guru dalam membuat tes.Subjek penelitian ini adalah calon guru di pendidikan bahasa inggris yang sudah mengambil mata kuliah Language Assessment.Untuk menjawab pertanyaan penelitian, penelitian ini melakukan interview kepada 5 partisipan mengenai beberapa pertanyaan seperti pengalaman dalam membuat tes, tantangan dan strategi dalam membuat tes.Hasil dari penelitian ini menunjukkan ada 5 tantangan yang dihadapi; mendesain rubric untuk soal essay, menulis soal dengan baik, membuat pilihan untuk pilihan ganda, menyeimbangkan skill dalam tes, dan mencari sumber.Dari tantangan yang disebutkan, mendesain rubric untuk soal essay merupakan tantangan yang sering dihadapi oleh calon guru. Penelitian ini juga menemukan bahwa menulis tes dengan melihat silabus, mendesain tes dengan jelas dan objektif, dan menulis rubric dengan jelas merupakan strategi calon guru dalam membuat soal. Ini merupakan strategi untuk membuat soal yang sesuai dengan level siswa dan juga mencapai tujuan tes.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

TABLE OF CONTENT

ADVISOR APPROVAL SHEET i

EXAMINER APROVAL SHEET ii

MOTTO iii

DEDICATION iv

ABSTRACT v

ABSTRAK vi

ACKNOWLEDGMENT vii

PERNYATAAN KEASLIAN TULISAN viii

LEMBAR PERNYATAAN PERSETUJUAN PUBLIKASI ix

TABLE OF CONTENT x

LIST OF FIGURE xii

LIST OF TABLE xiii

LIST OF APPENDICES xiv

LIST OF ABREVATION xv

CHAPTER I 1

INTRODUCTION 1

A. Research Background 1

B. Research Questions 5

C. Objectives of the Research 5

D. Significance of the Research 6

E. Scope and Limitation 6

F. Definition of Key Terms 7

CHAPTER II 9

REVIEW OF RELATED LITERATURE 9

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

A. Theoretical Framework 9
B. Previous Studies 22
CHAPTER III 26
RESEARCH METHOD 26
A. Research Design 26
B. Subject and Setting of Research 27
C. Data and Source of Data 27
D. Data Collection Technique 28
E. Research Instrument 29
F. Data Analysis Technique 30
CHAPTER IV 32
FINDING AND DISCUSSION 32
A. Research Findings 32
1. Challenges in Constructing Test Items 33
2. Strategies in Constructing Test Items 44
B. Research Discussions 49
1. Challenges in Constructing Test Items 49
2. Pre-Service Teachers' Strategies In Constructing Test Items 55
CHAPTER V 60
CONCLUSION AND SUGGESTION 60
A. Research Conclusion 60
B. Research Suggestion 62
REFERENCES 64
APPENDICES 66

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

LIST OF FIGURE

Figure 1. 1 Steps in Teaching Language 9

Figure 4. 1 example of stem which is not meaningful 52
Figure 4. 2 example of stem which is meaningful 52
Figure 4. 3 example of stem irrelevant material 53
Figure 4. 4 example of negative phrasing 53
Figure 4. 5 example of negative phrasing better 53
Figure 4. 6 example of interior blank 54
Figure 4. 7 example of partial question 54
Figure 4. 8 example of multiple choice items 54

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

LIST OF TABLE

Table 4. 1 Challenges in Constructing Test Items 35
Table 4. 2 Strategies In Constructing Test Items 45

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

LIST OF APPENDICES

Appendix 1 Interview Guideline for Pre-Service Teachers' Challenges and Strategies in Constructing Test Items 66
Appendix 2 Analysis of Pre-Service Teachers' Interview for Challenges and Strategies in Constructing Test Items 68
Appendix 3 Surat Tugas 105

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

LIST OF ABREVATION

1. LTA : Language Testing and Assessment
2. PPL : Program Pengalaman Lapangan

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

CHAPTER I

INTRODUCTION

In this chapter, the researcher explains the background of the study of this research, the research questions, the research objectives, the significance of the study, the scope and limitation, the definition of the key terms, and the previous study.Some points are displayed by the researcher to answer those headings.

A. Research Background

One component which teacher has to do for teaching is assessing students. Good assessment depends on the quality of test items. It means test items must be designed well. Assessment has been implemented by educators to measure learning progress and students' skill throughout the learning in life.Assessment is a process of collecting, analyzing and interpreting information about teaching and learning to make informed decisions that enhance student achievement and the success of educational programs[1]. Assessment is very important for trackings progress, planning next steps, and reporting students in learning[2]. Assessment is one of the important parts in education**.** The purpose of assessment is to give students opportunity to show what they have understood or to show what they have not understood.

Biggs stated that there are two major functions of assessment. The first is to show the learning process has succeed or not. The second is to clarify teachers'

[1](Rea-Dickens and Germanie, 1993; Genesee and Upshur, 1996; O'Mally and Valdez-Pierce, 1996)
[2]Dr.K.Anandan, *Assessment for learning* (Bharatidasan University), 2

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

expectations from students[3]. Gensee and Upshur argued that assessment involve comparison.[4]Taghi Jabbarifar found that there are five basic components in assessment, those are articulating the purpose of the educational system, identifying and collecting relevant information, having ideas that are valuable and useful to learners in their lives and professions, analyzing and interpreting information for learners, classroom management or classroom decision making[5]. It can be concluded that assessment is important because teacher can measure students' understanding. Moreover, teacher can perceive that the learning activity is successful or not.

An important element of assessment is the construction of test items. In the field of foreign language teaching, a test is a tool for measuring someone's ability, knowledge, and performance. It is a technique, procedures or items that require test takers' performance.[6]The teacher must understand who take the test. Besides, teachers also have to know test takers' background knowledge and have to make sure that the test is appropriate for the test takers. There was a research that demonstrated that test items' sentences can affect the psychological perspective that the students or examinees bring when considering a response. Anxiety, motivation, and performance are influenced by the sentences of test items and format[7]. There are some considerations in

[3]Biggs, J. 1998, "*Assessment and classroom learning: A role for summative assessment?*", Assessment in Education: Principles, Policy & Practice, Vol. 5, No. 1, pp. 103-110.

[4](Rea-Dickens and Germanie, 1993; Genesee and Upshur, 1996; O'Mally and Valdez-Pierce, 1996)

[5]Jabbarifar Taghi., "*The Importance of Classroom AssessmeRnt and Evaluation in Educational System*". Proceedings of the 2nd International Conference of Teaching and Learning, 2009

[6]H Douglas Brown,*Language Assessment: Principle and Classroom Practices*. (USA: Longman. 2004), 6.

[7]Board of Scientific Affairs of the American Psychological Association.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

item preparation that can influence its quality. Those are employing an appropriate item format, the level of words and determining the optimal number of response alternatives.

Language Testing and Assessment (LTA) course is obligated for pre-service teachers before they get microteaching class. It is important to be learned because in LTA class, pre-service teachers learn about some topics related with assessment. The topics learned in LTA class are the relation between teaching and testing, kinds of testing, validity and reliability in language testing, skills in designing test specification, understanding the concept of assessment in all skills. By the end of this course, students are expected to understand the fundamental concepts in language testing, understand approaches and techniques in constructing language testing and do a critical analysis on certain used tests.

Not only understanding topics about testing but pre-service teachers also have to understand other things must be done while teaching. Pre-service teachers must understand some things which are important before teaching, such as designing a lesson plan. Designing lesson plan here means setting the objectives, preparing materials, and choosing an appropriate strategy. During the learning, pre-service teachers can also give score from students' performance in the class. Pre-service teachers must also understand what must be evaluated from the materials and what must be tested to the students. That's why having an ability to construct good test items is important for pre-service teachers. Pre-service teachers also must have a strategy to construct good test

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

items. In fact, some pre-service teachers may face some challenges in constructing good test items.

The study conducted by Risalatil Umami focuses on identifying students' ability to construct test items based on the level of cognitive in bloom taxonomy. The finding shows that the most difficulties faced by participants is grammar with the total 84%.[8]Another study conducted by Dian Novita focuses onXX analyzing the reading comprehension questions on final test made by English teacher based on Barret's Taxonomy. The result is the level of barret presented on the final test 94 between HOTS and LOTS were not in balance number.[9] The next study which was revealed by Fadillah focuses on analyzing teacher's ability in designing assessment instrument [10]. The finding was from all of students teachers who design the assessment, 47% of them have been NB fulfill the requirement. It can be concluded that most of them can design the assessment based on standardization test.

However, the previous studies only focused on the difficulties in designing instrument and reading test faced by the participants. Most of the previous studies focused on the reading question. Besides, the participants of the previous studies were from students in university which has not got language

[8]Risalatil Umami, Thesis: "*Students Ability in Constructing Reading Question Items in Critical Reading Class".* (Surabaya: UINSA 2016)
[9]Dian Novita, Thesis: *"An Analysis of Reading Comprehension Questions Made By English Teacher In SMAN 2 Sidoarjo Based on Barret's Taxonomy"* (Surabaya: UINSA 2016)
[10]Nur Amalia Fadhila: *"Student Teachers' Ability in Designing Assessment Instrument at English Teacher Education Department UIN Sunan Ampel"* (Surabaya: UINSA 2015)

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

assessment course. While, the participants of the present research will be students in Language Testing and Assessment class

So, this research explored the strategies in constructing test item that used by the pre-service teachers and challenges faced by pre-service teachers in constructing test items. In the previous study, there is no researcher who explore about this case.By doing this research, it can inform the lecturer about the challenges faced by participants. So, the lecturer can make a revision about the method used to make students overcome their difficulties. Besides, it will give information about some strategies in constructing test items.

B. Research Questions

The research questions will be formulated as these following questions:

1. What challenges do pre-service teachers face in constructing test items?
2. What strategies do pre-service teachers use in constructing test items?

C. Objectives of the Research

The objectives of the research are below :

1. To explore pre-service teachers' challenges in constructing test items
2. To investigate pre-service teachers' strategies in constructing test items

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

D. Significance of the Research

The result of this research has expected to give a contribution for lecturer of English department, Education College Students, and further researcher.

1. For lecturer

   The result provides the challenges that faced by pre-service teachers in constructing test items. So, the lecturers can understand pre-service challenges in constructing test items. In that case, lecturers can tackle pre-service teachers' challenges.

2. For pre-service teachers

   This research provides information about the strategies in constructing test items which will be useful for pre-service teachers when they get teaching practice (PPL). On teaching practice, pre-service teachers are obligated to understand the ways or strategies in constructing test items.

3. For Further Research

   This research gives advantages in academic information for the future research in ELT research especially in challenges and strategies of constructing test items.

E. Scope and Limitation

There are two scope on this research, they are challenges and strategies. In the term of challenges, this research focuses on pre-service teachers' challenges while constructing test items. Besides, this research also focuses on strategies used by pre-service teachers when they construct test items. The limitation of

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

this research only focuses on the challenges and strategies in constructing test items. This research does not analyze the document of the final achievement test faced by pre-service teachers for final project of language testing and assessment class. Pre-service teachers here means fifth-semester students in the academic year 2017/2018. In language testing and assessment class, students are asked to construct final achievement test for junior high school or senior high school. However, this research limits to pre-service teachers who made final achievement test for senior high school students. In this research, the researcher used brown theory for analyzing challenges in constructing test items. Then, the researcher used harmer theory for analyzing strategies in constructing test items.

F. Definition of Key Terms

To have the same idea in this study, the researcher explains the terms used in this study, as the details are:

1. Test

According to Brown a test is a technique, procedures or items that requires performance on the test taker. There are four types of test, such as proficiency test, achievement test, diagnostic test, and placement test.[11].In this research, test refers to the final achievement test for senior high school

[11] H Douglas Brown,*Language Assessment: Principle and Classroom Practices*. (USA: Longman. 2004), 3.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

students made by pre-service English teachers for the final project. Test made by pre-service teachers are integrated skills.

2. **Challenge**

Challenge is a difficult task or something that is hard to do[12]. This research defined challenge as students' situation to do effort more when constructing final achievement test for the final project of Language Assessment class.

3. **Strategy**

Strategy is a method of approaching a problem that is planned designed for controlling and manipulating specific information[13]. In this research, strategies mean pre-service teachers' planned designed in constructing test items.

4. **Pre-service Teacher**

Pre-Service Teacher is students who are practiced on how to teach in appropriate way and the way to be an excellent teacher from their experience as students in the school[14]. Pre-service teachers in the research refers to students in fifth semester at Language Testing and Assessment class who have experience in constructing test items. In this class, pre-service teachers are asked to construct test items for the final project.

---

[12] Merriam-Webster, Inc, accessed on 29th of July 2020
[13]H Douglas Brown.Principle of Language Learning and Teaching. USA:Longman (2000)
[14] K. A. Sirotnik. *Reviewing School and Teacher Education: An Odyssey in Educational Change.* (Washington DC: AACTE, 2001)

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

CHAPTER II

REVIEW OF RELATED LITERATURE

This chapter discusses the theoretical framework in this research, and the researcher gives some previous studies in order to give clear differentiation among the existing research.

A. Theoretical Framework

In this part, the researcher provides some theories related to the research. The theories are about the relation between teaching and testing, types of tests, criteria of a good test, challenges in constructing test items, and strategies in constructing test items.

1. Teaching and Testing

Teaching is an activity to help someone do something, give instruction, and understand the study[15]. In teaching Language, there are ten processes which are understanding students' needs, having goal and objectives, test tasks, Language and skills, sequence, materials, teaching, reflection, evaluation, and revisions[16]. It will be formulated with a picture below.

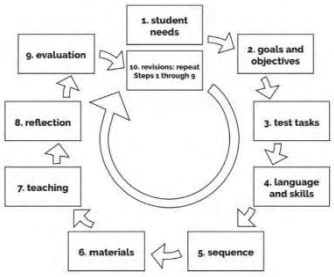

Figure 1. 1 Steps in Teaching Language

[15] H Douglas Brown, *Teaching by principle. An Interactive Approach to Language Pedagogy*. (USA: Longman. 2001), 55

[16]

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

From the figure, it can be seen that students' need is the first process. Here, the teacher should identify what students need in the classroom. To get what students need, the teacher can do two things such as observation in the classroom and asking students feedback through informal conversations. Secondly, the teacher must create goals and objectives in each lesson based on students' needs. According to Graves, goals mean the purpose of teaching while objectives are goals into statements that can be learned and measured[17]. The third is the test tasks. In this part, teacher must be creative in designing English tasks to measure students' performance for specific objectives such as a project or presentation. The next is language and skills. Here, teacher must guide the test while other teachers record the explicit Language used to complete the task. Then, teacher must prepare the selected Language and skills into a sequence. Teacher must consider the time while teaching. The next is materials for teaching. The lesson objectives can be used to create students' activities, homework, and quiz. It will be helpful for students to understand Language and skills and to succeed in the unit.

The next step is teaching the unit. This part is important for teachers because teachers not only conduct the lesson but also look carefully at students' reactions, behavior, and performance in relation to the lesson's objectives. Reflection is the next process of teaching. Teachers return to

[17]Graves K, *Designing language courses: A guide for teachers.*(New York: Heinle and Heinle, 2000), 58.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

their notes and make observations based on their experience. Next, evaluation is also one of parts in teaching. Teachers evaluate the unit and give advice for the next round of revisions. To evaluate the lesson, asking students' feedback by online surveys can be used as the tool. Besides, teachers and students can do conversations informally to know what must be evaluated. The last is revisions. It is actually a return to the first step in the process. In this point, teachers will discuss the revisions and formulate strategies on how to be best in teaching.

As mentioned before, evaluation, reflection, and assessment are important parts in teaching. As stated by Brown, teaching and evaluation are partners. Evaluation can be done in some ways such having a test. As students in the education department or usually called as pre-service teachers, understanding the way to construct good test items is obligated. It is because pre-service teachers will be real teachers in the future and will teach in real situations. Teachers will spend a lot of time to test, assess, and evaluate students. Assessment has many functions; it is to measure students' ability at the end of the course, it is to measure students' progress, place students into different stages, and others. To assess students, teachers or instructors must have good test items. Teacher must clearly understand the purpose of the test before giving test to students. The test's purpose also influences test design, item development, and how students' performance will be reported.[18] There was a research state that anxiety, motivation, and

[18]Michael Bay Borelli, *Considerations For Developing Test Specifications*

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

performance are affected by the sentences of test items and format[19]. It means that when teachers assess students, they must have good test items to get good results.

2. Types of Test

Test is usually used in educational program. Test is a way to measure someone's ability, knowledge, and performance. There are some types of test. Some tests measure general ability while others focus on specific objectives. According to Hughes, there are four types of test, they are proficiency tests, diagnostic tests, placement test, and achievement tests. All of these types of test have their different purpose and functions. These tests are commonly used by educators to assess students[20].

a. Proficiency Tests

Proficiency tests are tests to measure people's ability in language regardless of any preparing they have had in that Language[21]. It gives general information about student's knowledge and ability. The contents is what test takers have to be able to do in the Language. The function of this test is to show whether the test takers have reached the standard or not[22]. This test also as the standard test to understand people's ability in all Language skills. In many countries, this test is usually used to find an

For Common Core Assessments, (Pearson, 2010) 12
[19]Board of Scientific Affairs of the American Psychological Association
[20]Arthur Hughes,*Testing for Language Teacher* (New York: Cambridge University Press,2007), 9
[21]Ibid
[22] Ibid

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

international job or receive a certificate. Mostly, the university uses proficiency test as the requirement for graduation.

b. Diagnostic Tests

According to Harmer, diagnostic tests are tests that used to reveal learner's difficulties and gap in their knowledge during the learning process[23]. This test is also used to identify students' weakness and strengths[24]. One of the simplest diagnostic test is pretest. Pretest is usually done at the beginning of the lesson. By having this test, teachers can be easier to determine what strategies will be used in learning process and what materials will be given more to the students. Usually, after having pretest at the beginning, some educators also have post test. The function is to understand students' progress after having a lesson.

c. Placement Tests

Placement tests are tests that usually used by school and university to place students into some levels. Hughes argues that placement tests will help students to get the most appropriate program based on students' abilities[25]. Students who have the same ability will be in the same class. The advantage of this test is teachers can choose appropriate materials for each class. For example, in the low level, teachers may prepare the easiest materials to make them understand the materials. Then, in the high level, teachers may have some difficult materials.

[23]Harmer J, *The practice of English language teaching* (Harlow: Pearson Education Limited,2004), 380
[24]Ibid
[25]Ibid

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Besides, it is also useful for the teacher because teachers can choose an appropriate method based on the level.

d. Achievement Tests

Achievement tests are designed to measure students' language skill in relation to the syllabus[26]. According to Hughes, there are two kinds of test; final achievement test and progress achievement test. Further, each definition will be discussed more specifically below.

1) Progress achievement test

According to Hughes, progress achievement tests are planned to measure students' progress[27]. The content of this test is only assessing materials that have been taught by teacher. This test can be a reflection for the teacher whether students have understood the material or not. This test is usually given to students every few weeks to know how well students understand the materials that have been taught by teacher.

2) Final achievement test

Final achievement test is conducted at the end of a course[28]. Some testers argue that the final achievement test's content must be based on syllabus, book, and other materials used. It is usually called as syllabus-content approach. The disadvantage of this test is that if

[26]Harmer J, *The practice of English language teaching* (Harlow: Pearson Education Limited,2004), 380
[27]Arthur Hughes,*Testing for Language Teacher* (New York: Cambridge University Press,2007), 13
[28]Arthur Hughes,*Testing for Language Teacher* (New York: Cambridge University Press,2007), 10-11

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

the syllabus is not good enough or other materials are badly chosen, the test result can be bad.

In Indonesia context, final achievement tests or usually called *"ulangan akhir semester"* are mostly used by teachers. In final achievement test, teachers usually use both multiple choice and short essay. Before having final achievement test, teachers must make sure that they have taught all materials. In having final achievement test, teachers must have good syllabus. As mentioned before, if the syllabus is not good enough, it will influence the test results. While progress achievement test or usually called *"ulangan harian"* only to evaluate in the middle of the lesson. This test is usually done after teachers explain some materials. This test is to know students' progress in the lesson. The difference between these two tests is that the final achievement test assesses all lessons in one semester while the progress achievement test assesses some materials that teachers have been taught.

3. Constructing Test Item

As students in education department or usually called as pre-service teachers, constructing test items is obligated. Pre-service teachers are obligated to understand the way to construct good test items. Understanding the way to construct good test items is important because as pre-service teachers, they will construct many test items for assessing students. Criteria for good test items are difficult to state in absolute terms. It is difficult to

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

determine that test items is "good" or "bad". It is also difficult to establish criteria for good test items for a number of reasons. One reason that makes constructing test item so difficult is technical skills. Each test item has different criteria for good test. There are some criteria of good test items stated by experts.

According to Steven J Osterlind there are some criteria of good test items. The first criterion is there must be a high degree of coherence between a particular item and the key objective of the total test. The second criterion is the objectives must be defined clearly. In constructing test items, testers must understand what the objective of the test is. The third criterion is the each test item's contribution to measurement error in a test's scores should be minimized. The fourth criterion is the format of the test must be suitable to the goals of the test. The fifth criterion is the tests are well written. It should follow the standards such as grammar, diction, spelling, punctuation, usage, and syntax. The sixth is legal and proper questions. It does not mean that all test items must be originally constructed[29]

4. Challenges in Constructing Test Items

Challenge is a difficult task or something that is hard to do[30].Constructing test items is complicated. It must be practiced so many times to be mastered in constructing test items. As pre-service teachers, they may face some challenges in constructing test items. It happens because

[29]Steven J Osterlind, *Constructing Test Items: Multiple-choice, Constructed-Response, Performance, and Other Formats* (New York: Kluwer Academic Publishers,2002), 40
[30]Merriam-Webster, Inc, accessed on 29th of July 2020

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

there are some considerations before constructing test items such as understanding students' level. Besides, it is challenging because there are some steps must be done when constructing test items. According to Brown, there are some steps to construct good test items those are; test toward clear, unambiguous objectives, draw up test specifications, draft the test, revise the test, final-edit and type the test, utilize the feedback after giving the test, work for washback[31].The task of constructing good test items is difficult because writing correctly and briefly is challenging. Besides, the writer must be correct in the grammar, punctuation, spelling, and syntax when constructing test items[32]. Other challenges is the writer should be creative when constructing test items. Recent research has revealed that the sentences and the format of test items can influence examinees' when answering the test. As revealed by Board of Scientific Affairs of the American Psychological Association, the sentences and format of test items can affect examinees' anxiety, motivation, and performance[33].

Sam Warib states that most students are unwilling to construct test items because of a lack of grammar and vocabulary[34]. Those components must be mastered by examiners in constructing test items. The following concept are explained below:

a. Vocabulary

[31]H Douglas Brown,*Language Assessment: Principle and Classroom Practices*. (USA: Longman. 2004), 411
[32]Steven J Osterlind, *Constructing Test Items: Multiple-choice, Constructed-Response, Performance, and Other Formats* (New York: Kluwer Academic Publishers,2002), 26-30
[33]Board of Scientific Affairs of the American Psychological Association
[34]Mun.Fika et al, *Complete English Grammar (*Surabaya: Apollo, 1991)

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

According to Aitchitson, vocabulary is a set of words known to a person that is part of a specific language. It seems impossible to learn a Language without having vocabularies to support. Vocabulary takes an important part in language in which the vocabulary will make a Language meaningful. Vocabulary is an essential part in Language, to master a Language automatically have to master vocabulary. According to the expert, the basic of mastering vocabulary is learning how to determine the meaning[35]. Most of pre-service teachers have problems in writing test because of lacking vocabulary. Pre-service teachers feel difficulty in choosing the word which is suitable with the subject. As stated by Sam Soesanto good writing a test is made from words which is suitable with the content.

b. Grammar

Grammar is important for pre-service teachers in constructing test items. Grammar is one of the English skills that must be mastered by pre-service teachers. By mastering grammar, pre-service teachers can produce correct sentences. Sam warib states that mastering grammar is good for listening, speaking, reading, and writing[36]. Good writing usually comes from good grammar. Based on these statements, it is very important to mastery grammar in order to be able to construct good sentences for test items. Without having grammar knowledge, pre-

[35]Mun.Fika et al, *Complete English Grammar (*Surabaya: Apollo, 1991)
[36]Mun.Fika et al, *Complete English Grammar (*Surabaya: Apollo, 1991)

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

service teachers will not be able to construct good test items. So, pre-service teachers must be able to construct good test items with correct grammar.

5. Strategies in Constructing Test Items

Strategy is a method of approaching a problem that is planned for controlling and getting an information[37].Constructing test items is one of skills that must be had by pre-service teachers. By understanding the strategies to construct test items, it will be easier for pre-service teachers to test students. In constructing test items, pre-service teachers may use some steps and preparation which can be different with others. Each pre-service teacher may have their own strategies. As stated by an expert, there are some steps in constructing test items those are; assess test situation, decide what will be tested, balance the elements, weight the scores, and make the test work[38]. It will be explained below :

a. Assess test situation

To construct test items, pre-service teachers is not only preparing the questions but also understanding the situation which test takes place. Understanding the situation here means pre-service teachers must understand several things, those are; how much time will be given to students while doing the test, when and where students will take the test, and how much time pre-service teachers need for giving mark.

[37]H Douglas Brown.Principle of Language Learning and Teaching. USA:Longman (2000)
[38]Ibid

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

b. Decide what will be tested

In this part, pre-service teachers must have a list about what will be included in the test. The test must be a representative sample from whole list . To make a list, understanding the syllabus items is needed. By understanding syllabus closely, pre-service teachers can determine what kinds of materials or topics are appropriate for students. Not only determining the materials but pre-service teachers also determining what skills will be tested for students. It means, pre-service teachers take a conscious decision to include or exclude skills such as writing or reading comprehension. By including a representative sample from the whole list, students' success or failure with those test items will be a good indicator of how well students have understood all of the materials they have studied.

c. Balance the elements

Pre-service teachers take an important part when constructing test items. They must consider some things such as including direct or indirect test items. In constructing test items, the elements must be balanced. Elements here means how many multiple choice test and how many essay will be included. Not only that, but the skills also must be balance. Pre-service teachers must consider how many numbers of each skills. Balancing the elements can estimate how long time must be given in each section of the test.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

d. Weight the scores

Score is also important in constructing test items. Teachers' perception of students' success or failure depends on how many marks will be given in each section of the test. For example, the score of multiple-choice questions will be different with short essay. In giving the marks, teachers must also consider about level of difficulty in each questions.

e. Make the test work

Pre-service teachers have to try out some items of the test or whole tests to other students which have the same level with the real candidates before giving the test to real candidates. Frequently, other students have problems which pre-service teachers are not aware. Besides, they also come up with possible answers and alternatives which pre-service teachers had not anticipated. After, trying out to other students, pre-service teachers can make revision based on other students' reaction.

In constructing test items, those elements should be conducted because those elements are related. First, assess test situation means pre-service teachers consider several things such as how much time will be given, when students will take place, and others. Second, pre-service teachers also have to decide what will be included in the test. Before designing the test, pre-service teachers should understand the syllabus clearly. Besides, pre-service teachers also should have a list about kinds of materials and skill in the test. Third, pre-service teachers should

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

balance the elements. Elements refer to how many multiple-choice and essay questions will be included in the test. After designing the test, pre-service teachers also consider about the score for each questions. However, before giving test to students, pre-service teachers should try out some items to other students who have same level. From these steps, assess test situation is the most important because it is the beginning before designing the test.

B. Previous Studies

There are some research that have a similar topic with this research which conducted by other researchers. Moreover, the researcher provides the previous studies below in order to know the differences among upcoming research. The first study was done by Risalatil Umami.[39] This study explains the ability of students in constructing test items based on the cognitive level in bloom taxonomy's perspective and the challenges faced by students when students construct the test items. This research found that grammar is the most difficult things with the total 84%. The other difficulties students' faced were constructing a sentence in English language correctly and accurately, spelling, punctuating, finding appropriate vocabulary and making clear questions.

It is important to focus on understanding students' ability in constructing test items. Other researcher also conducts about students' ability. This study

[39]Risalatil Umami, Thesis: "*Students Ability in Constructing Reading Question Items in Critical Reading Class"*. (Surabaya: UINSA 2016)

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

was conducted by Fadhilah.[40] Her research has the same goal about knowing students' ability in constructing test items. The different is the subject and the measurement. Fadhilah measured students' ability based on the requirement of standardization assessment and also identify the factor that make unable to fulfill the assessment requirement. The participants of this research was 19 students from PPL class of English Teacher Education Department at State Islamic University of Sunan Ampel Surabaya. The finding was from all of students teachers who design the assessment, 47% of them have been fulfill the requirement, 37% of them could not meet the indicators, and 16% of them could not meet two of the indicators.. Then, the most causing factor of the weakness in their assessment instrument is the detail with the total 31%. So, it can be concluded that most of them can design the assessment based on standardization test.

There are some other previous studies which analyze about questions based on its criteria. The first study was conducted by Dian Novita.[41]Her study observed the reading questions of final test made by English teacher reflect criteria of Barret's Taxonomy or not and in what level dominant found it. It was aimed to classify into good reading comprehension questions. The result of her research is the level of barret presented on the final test 94 between HOTS and LOTS were not in balance number. The second study was done by Nur

[40]Nur Amalia Fadhila, Thesis: *"Student Teachers' Ability in Designing Assessment Instrument at English Teacher Education Department UIN Sunan Ampel"* (Surabaya: UINSA 2015)
[41]Dian Novita, Thesis: *"An Analysis of Reading Comprehension Questions Made By English Teacher In SMAN 2 Sidoarjo Based on Barret's Taxonomy"* (Surabaya: UINSA, 2016).

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Sartika.[42]Her study observed on the quality of test items used in semester examination. The purposes of this research were to decide whether the quality of English semester test items fulfilled the criteria of a good test or not. The results were determined that 10 poor items, 24 satisfactory items, 12 good items, 3 negative items, 1 excellent items. While according to the level of difficulty were determined that 17 easy items, 16 difficult items, and 17 average items. The third study was revealed by Arinal Rahman.[43] This study focussed on describing the quality of test items in final examination at PPB based on validity, reliability, item difficulty, item discrimination, and distractor efficiency. The researcher used document analysis technique to analyze the data. The results showed that the test items in final examination at PPB were very good.

The next previous studies is about analyzing questions made by teacher. There are two previous studies discussed about it. The first was done by Zuliati Rohmah and the second was done by Yumna Shaumu. The study by Zuliati Rohmah was aimed to discuss English teacher's questions in reading class at MAN III Malang. This is non participant observation. The researcher collected the data by the researcher itself and also field notes. The result shows that the questions that distributed by teacher did not allow students to talk longer and have critical thinking. The study by Yumna Shaumu was discussed about analyzing the constructions of multiple-choice items, particularly the multiple-

[42]Nur Sartika, Thesis: *"An Analysis of Test Items Based on the Criteria of Good Tests. (Bandar Lampung: Lampung University, 2015).*
[43]Arinal Rahman, Thesis: "*An analysis of test items quality in English final examination at Language Development Center"*(Banjarmasin: IAIN Antasari 2016)

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

choice construction includes 3 categories, those are: format concerns, the writing of the stem, and the writing of the options. This research focused on analyzing the multiple-choice format, stem, and options by using cr
iteria of validated multiple-choice items construction.The first category is that 50 multiple-choice items have good format-construction. The second category is about the item-construction. The last category is about option construction.[44]

Generally, the study of test items discussed analyzing the test based on some criteria, analyzing students' ability to construct reading test items, and analyzing tests made by teachers in school. Therefore, this research focus on challenges and strategies faced by Pre-Service Teachers in constructing test items.

[44]Yumnaa Shaumu, Thesis: "*An Analysis of Multiple Choice Items of Computer Based English Test at Smp Alkhairiyah Surabaya*" (Surabaya: UINSA 2019)

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

CHAPTER III

RESEARCH METHOD

In this chapter of this research, the researcher explains the research method of this research and how the researcher collects the data. The detail will be explained below.

A. Research Design

Research design is a research plan will be conducted by researcher[45]. This research used a qualitative method to explore some phenomena by interviewing participants. According to Creswell the aims of a qualitative method is to explore a detailed understanding of central phenomenon and subject of qualitative research is a small number of individuals so that the participants' views are obtained[46]. This research used a qualitative method because it was described as phenomenon about pre-service teachers' challenges and strategies in constructing test items.This research investigated pre-service teachers' experiences in constructing test items, pre-service teachers' challenges in constructing test items, and strategies used by pre-service teachers in constructing test items.

[45]Suharsini Arikunto, *Prosedur Penelitian Suatu Pendekatan Praktek*(Jakarta : Rineka Cipta, 2002),51
[46]John W Creswell, *Educational Research Planning* (UK Sage Publication, 2012), 17

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

B. Subject and Setting of Research

The subject of this research was five Pre-Service English Teachers in State Islamic University of Sunan Ampel Surabaya who elected with the criteria (1) have taken Language Testing and Assessment class (2) have experience in constructing final achievement test for senior high school students. The researcher interviewed five pre-service teachers who made final achievement test for senior high school students. The researcher chose test for senior high school because the level of the questions is higher than junior high school. The researcher chose these five participants by having preliminary research. The researcher asked about a list of pre-service teachers who made final achievement tests for senior high school in the final test of Language Assessment Course. This research used these five participants because they had experience designing final achievement tests for senior high school while other students made final achievement test for junior high school.This research took place in two classes of Language Testing and Assessment class, which taught by the same lecturers. In these classes, pre-service teachers had experience in constructing test items for the final project of Language Testing and Assessment class.

C. Data and Source of Data

In this part, the researcher provides data used by the researcher and source of data used by the researcher.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

1. **Data**

The data used by researcher was pre-service teachers' challenges and strategies in constructing test items. Those are pre-service teachers' responses about experience in constructing test items, challenges in constructing test items, and strategies used in constructing test items.

2. **Source of Data**

In this research, the sources of data were from pre-service teachers' in English Education Department at UIN Sunan Ampel Surabaya. The data needed in this research was the recording from interview about pre-service teachers' experience in constructing test items, challenges faced by pre-service teachers in constructing test items, and strategies used by pre-service teachers in constructing test items.

D. Data Collection Technique

According to Creswell, there are some techniques to get the data such collecting data through observations, interviews, documents, and visual materials, and it is better to provide recording information[47]. In this research, interview is used as the main tool for collecting the data. The interview used in this research was semi-structured interview.The researcher interviewed these participants on April. The online interview has been used by researcher to collect the data. The researcher interviewed these 5 participants in the same day

[47]John W Creswell. Research Design Qualitative , Quantitative and Mixed Methods Approaches (UK Sage Publication, 2009), 178

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

by phone. It was only once because the deep information had got by researcher. The researcher interviewed pre-service teachers to reveal challenges faced by them in constructing test items and strategies used by them in constructing test items.

E. Research Instrument

To answer this research, the data was collected through interviews. As stated by Uwe Flick, interviews are divided into three types, those are unstructured interview, a semi-structured interview, and structured interview[48]. Based on that statement, this research used semi-structured question. The researcher had an interview guideline, but it can be modified related to the interviewees' answers. The interview guideline was required to help the researcher in the process of interviewing. There were some general questions in interview guideline:

1. Pre-service teachers' experience in Language Testing and Assessment class. It means pre-service teachers knowledge about Language testing such as knowledge about how is good test, how to construct good tes, and many more about assessment. This knowledge could be got from Language Testing and Assessment Class.
2. Pre-service teachers' experience in constructing final achievement test for senior high school. It refers to pre-service teachers' experience when

[48]Uwe Flick, *An Introduction to Qualitative Research*, 4th ed. (Los Angeles: Sage Publications, 2009), 223.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

designing tests for their final achievement test of Language Testing and Assessment Class. However, not all pre-service teachers had this experience because most of them designed test for junior high school students.

3. Pre-service teachers' challenges faced in constructing final achievement test. It refers to their challenges in constructing test for the final project of Language Testing and Assessment Class.
4. Pre-service teachers' strategies used in constructing final achievement test. This point means pre-service teachers' steps in constructing test.

F. Data Analysis Technique

The researcher analyzed the data using descriptive qualitative descriptive method. As stated by Creswell there are six steps for analyzing data which are collecting the data, organizing and preparing the data, reading all the data, coding the data, coding to build description or themes, and interpreting the findings[49]. Firstly, collecting the data, the data were collected by interviewing participants. Second, the collected data were prepared and organized before going to the next step. The data were arranged into different types depending on the source of the data. The next step was reading all the data in order to get information as much as possible from each category. Then, the interview transcript coded by researcher. This step also identified which information will be used or need to be reduced. By doing this, it selected specific information

[49]John W Creswell, *Educational Research Planning* (UK Sage Publication, 2012), 17

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

that focus on the study. After having some categories based on the interview transcript, the next step was achieving the category in detail. The researcher analyzed the data using Brown theory. The next step was interpreting the findings. The analyzed data explained in detail in the finding and discussion. The last, the researcher explained the conclusion

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

CHAPTER IV

FINDING AND DISCUSSION

This chapter presents the finding and discussion of the research. The findings answer the two research questions, namely: challenges in construction test items and teachers' strategies in encountering the challenges. The organization of the findings is arranged according to the two research questions. Then, it discusses the main findings and link them to the relevant theories and previous studies. The detail information of the findings and discussion is explained in the following sections:

A. Research Findings

To obtain the data about pre-service teachers' challenges and strategies in constructing test items, this research interviewed 5 students who took language and assessment course. The researcher gathered the data from 5thApril-20thApril 2020. The result of interview analysis indicated that pre-service teachers had some challenges in constructing test items. The challenges are categorized into three aspects: knowledge and experience, language, and sources. The data also indicated that pre-service teachers used different ways to overcome the challenges. The detail information of the challenges and strategies used in constructing test items is described in the following section:

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

1. Challenges in Constructing Test Items

The result indicated that all pre-service teachers were aware constructing test items were not easy. They found some challenges in doing it. Each participant had their own criteria on how to construct the test. For example, participant 1 argued that a good test should assess students' background knowledge which covered all skills: speaking, listening, reading, and writing. Similarly, participant 5 also said that a good test should consist of the materials that students have learnt and the items or questions in the test should match to the students' level. Importantly, the instructions of the test should be clear. This statement indicates that by giving clear instruction, these pre-service teachers may expect students will understand how to answer the questions. Another important criteria for good test identified was providing with a clear rubric.

"Test which appropriate with materials have been taught by teacher. Besides, good test also has to include all skills in English" (Participant 1)

"Test which appropriate with the level. It also includes all materials that students' got and has clear instruction. The level of difficulty is also important. So, the test is not too easy or too difficult" (Participant 5)

"Good test depends on rubric made by teachers. It also depends on syllabus and has appropriate topic for senior high school" (Participant 3)

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

To describe the challenges in constructing test items, the researcher interviewed 5 pre-service teachers who took language assessment course. The results of interview indicated that pre- service teachers encountered several challenges when constructing test items for the final project of language assessment course. The challenges are categorized in three aspects: knowledge and experience, language, and source.

The first aspect is knowledge and experience. Pre-service teachers defined knowledge and experience as background knowledge in constructing test items. Background knowledge means the level of knowledge the participants who will complete the test. Experience refers to participants' experience in designing test before taking language assessment course. The second aspect is language. Language refers to participants' ability in writing good sentences for test items. This aspect is important as when pre-service teachers were able to construct good sentences, students could understand the test easily. The third aspect is resource. It relates to ability in finding appropriate resources for constructing test items.

The challenges are divided into two categorize such as challenges in multiple choice questions and challenges in essay questions. The results of

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

interview shows pre-service teachers encountered different challenges in constructing test items, namely: making the rubric, covering all skills, deciding the level of difficulty, constructing good questions, and finding references, and etc. Table 4.1 summarizethe challenges of pre-service teachers in constructing test items.

Table 4. 1 Challenges in Constructing Test Items

Aspects	Challenges in Multiple Choice Questions	Challenges in Essay
Knowledge	Covering all skills in the test	Making rubric for open-ended questions
	Deciding the level of difficulty	
	Designing test creatively	
Experience	Constructing good questions	
	Writing the options for multiple choice questions	
Source	Finding references for the test	

Table 4.1 was from pre-service teachers' interview data. As stated in the table, there are some aspects of challenge faced by pre-service teachers in constructing test namely: knowledge and experience, language, and source.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

To gain the data about pre-service teachers' challenges, the researcher did interview sections with pre-service teachers. There are 17 items for the interview questions. The type of interview questions is open-ended questions *(see Appendix 1)*. The result of pre-service teachers' interview related to challenges is explained below completed with pre-service teachers' utterances to support the data.

a. Knowledge and Experience

Pre-service teachers' challenges in knowledge and experience is defined as their challenges in four points, those are: making rubric for open-ended questions, including all skills in the test, deciding the level of difficulty, and designing test creatively. Moreover, it is also related with their experience in assessment area. Some of them stated that they had experience in designing test but they did not know how to design test correctly. This statement indicated that some participants construct test item without understanding about assessment knowledge. The detail information will be explained below.

The first challenges in knowledge and experience is making rubric for open-ended questions. This is the most common challenges faced by participants. Participant 4 argued that information about

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

designing rubric had been taught by lecturer but it was confusing to be applied. Besides, some of them stated that it was their first time to design rubric based on the knowledge given by lecturer. Moreover, other participants were confused how to determine the criteria and how to assess the test objectively.

"Mmm.. for essay, the challenge is looking for appropriate topic and also making the rubric. Actually my lecturer had taught me but maybe because it is my first time in designing rubric based on the knowledge, so it is a big challenge for me"(Participant 4)

"At that time, we did not only design the test but also make the rubric. The rubric for multiple choice is easy enough but rubric for essay (open-ended question) is so difficult because we are not only including the answer but also including the criteria. Writing the criteria and the score for the test is difficult,"(Participant 5)

"Designing open-ended questions is not a challenge for me but designing the rubric for open-ended questions is challenging. It is because we have to have criteria for each number." (Participant 1)

"For me, assessing writing is difficult enough especially in making the rubric because we have to assess students objectively. It's difficult enough because sometimes I still design the rubric subjectively." (Participant 2)

"There is no challenge in multiple choice questions but the difficult thing is designing rubric for open-ended question" (Participant 3)

Dealing with the challenges, pre-service teachers had different ways to overcome it. Most of them overcome it by looking references

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

from internet and adapting it. While, participant 4 said that designing open-ended question as simple as possible will make easier in designing the rubric.

"I adapt it from existing rubric"(Participant 1)

"I search in internet" (Participant 2)

"I look some open-ended rubrics in internet with the same topic. Then, I see what are assessed in that rubric. Then, I try to design rubric"(Participant 5)

"I adapt rubric from the assignment had given by my lecturer. I choose one of those rubric"(Participant 3)

"Mmm.. design test as easy as possible hehe. I mean like no ambiguous rubric. I also have a trick. So, for example the score for one question is 1 until 4, if student have some errors in vocabulary the score can be 3 or have error in grammar and vocabulary the score can be 2" (Participant 4)

The second challenges in knowledge and experience aspect is covering all skills in the test. The result revealed that all participants got difficulties in covering all skills to the test. Therefore, most of references in internet only include reading and writing questions. For example, participant 1 argued that including all skill with balance was difficult enough because there was no list of material when constructing test items. Whereas, participant 3 stated that most of

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

references from internet or book only included reading and writing questions. Similarly, participant 2 claimed that the test only included reading and writing skill. Those statements indicated that designing reading and writing questions are easier than others because there are so many references about that.

"For balancing all skills, I just use my feeling. Then, I think balancing all skills is difficult because I didn't make list of material before designing the test so it's difficult enough" (Participant 1)

"Actually the lecturer asks us to include all skills with balance, but my test is not balance hehe. So many reading questions because I look from my references and it also has many reading questions"(Participant 2)

"It is difficult because I look from book and it's all only about reading and writing" (Participant 3)

The third challenges in knowledge and experience aspect is deciding the level of difficulty. The result indicated that pre-service teachers were aware deciding the level of test for senior high school were not easy. It was difficult enough to determine whether the question is too easy or too difficult for students. It is supported by participant 3 who said that looking for essay questions for senior high school is hard enough. Similarly, participant 5 argued that deciding level for senior high school is difficult enough. It can be said that

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

looking for test which appropriate for senior high school level is not easy. It is because there are several things to be considered such as the level of difficulty.

" I get difficulty in looking questions for essay and multiple choice which is not too easy or too difficult for senior high school. Because I'm afraid that my test is too easy or too difficult for senior high school student"(Participant 3)

"Determining essay test and difficult test for senior high school is not easy." (Participant 5)

The fourth challenges in knowledge and experience aspect is designing test creatively. Creative is needed to design the test because it will not make students bored with the test. Some pre-service teachers said that they got this challenges because they have not experienced yet in constructing test items. There are two participants who felt this challenge. First, participant 4 stated that designing test creatively was also a challenge. The challenge was in determining kind of questions for the test. In addition, participant 5 also claimed that designing test with several variation of questions was hard enough.

"Mmm.. it's not too difficult but it's not easy" (Participant 4)

"Yes, it's difficult enough. Because when I design the test I design it suddenly without have a list what kind of questions will be in my test. For example, I have used this kind of question then I will be confused what's next" (Participant 5)

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

b. Language

Pre-service teachers' challenges in language is defined as their challenges in understanding the way to write the test. Language is important in writing test because it can affect students' understanding in the test. In language aspect, the researcher found two points, those are; constructing good questions and writing the options for multiple choice questions.

In the finding, there are 4 participants who faced the challenges in the area of language. Two participants had challenge in constructing good questions while two others had challenge in writing the options for multiple choice questions. Participant 5 felt afraid that test made by her can not be understood by students because the sentences is not easy to be understood. While participant 1 stated that re-read several times the test before it's submitted is important.

Meanwhile, two others participant who faced challenge about writing the options for multiple choice stated that writing the options was also challenging because there are some considerations in writing the choice. Participant 2 and 3 had same idea about it. They stated the

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

level of the word and the length of sentences must be same. So, those are some factors that make them feel difficult.

"Constructing the question with good sentence is hard enough because I never make it before. Even I feel that the grammar is correct but I'm afraid that my test is not as the requirement of good test" (Participant 5)

"Writing good sentence actually is not a big challenge but re-read the sentence before giving to students is also important. It is to make sure that my test is good enough" (Participant 1)

"I also have difficulties in writing the choice for multiple choice test. As I remember, my lecturer told me that the level of the word and length of the sentence in the choice must be same." (Participant 2)

"Mmm.. at first, I feel that writing the choice for multiple choice is difficult. So, I try to remember the example and explanation from my lecturer" (Participant 3)

Those four participants had their own challenges dealing with language aspect. Besides, they also had their own way in encountering the challenges. Dealing with challenges in writing the choice,participant 2 overcame it by writing all choice based on the test. Then, participant 3 solved it by looking for example from lecturer's explanation to be adapted. Meanwhile others participants who had challenge in writing the sentences for the test overcame it by re-read test made by them. It is

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

to make sure that their test is easy to be understood and there is no grammatical error.

"I have a trick. So, I read the question then I have a list of the choice which have the same level of the word and the choice is from the text. Then, I also concern in the length of the word"(Participant 2)

"At first, I feel it's so difficult to make the choice but I look at example which was given by my lecturer."(Participant 4)

"That's also challenge that I faced but actually it's not a big challenge. I just re-read before submitting it. (Participant 1)

"Yes, I get difficulty in writing the sentence because I have not experienced yet in writing test. In my opinion, even my grammar is correct but it doesn't mean the test is easy to be understood" (Participant 5)

c. Source

This aspect is defined as participants' challenges in looking for the material and example for the test. Pre-service teachers were confused to get references for listening and speaking test. They argued that most of references only included reading and writing skill in the test. Meanwhile, the lecturer asked pre-service teachers to include all skills in English for the test.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

The data from interview showed that most of pre-service teachers used internet and book for the references. There were two pre-service teachers who got challenge in finding the material for listening and speaking skill. For example, participant 1 stated that listening was the most difficult because the audio was difficult to be searched. In other hand, participant 2 said that designing speaking test is the most difficult. In short, both of them do not include speaking or listening skill in the test because they get difficulties in finding the references.

"Designing speaking test is difficult but I didn't include speaking skill because it's too complicated. I also didn't find references that include speaking test" (Participant 2)

"Listening is difficult because I have to look for the audio. It's impossible if I make the audio by myself because my speaking is not good enough"(Participant 1)

2. Strategies in Constructing Test Items

After finding pre-service teachers' challenges in constructing test items, the researcher focuses on the second research question since the first research question is pre-service teachers' challenges in constructing test items. While the second research question focuses on the strategies in constructing test items. To get the data, the researcher conducted an

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

interview with 5 pre-service teachers from 2 classes of language assessment class which is taught by the same lecturer.

According to the interview, the researcher found the strategies that used by pre-service teachers in constructing test items. The strategies used are shown in the table :

Table 4. 2 Strategies In Constructing Test Items

Participant 1(P1)	**Participant 2(P2)**
a. Check the material b. Search some references c. Try to write the test d. Re-read the test	a. Outlining b. Look for some references c. Design the test d. Make the rubric
Participant 3(P3)	**Participant 4(P4)**
a. Make the rubric b. Check the material c. Search some references d. Make the answer key	a. Read the syllabus b. Make outline c. Read references d. Start to design test with clear instruction
Participant 5(P5)	
a. Look at list of materials b. Look for example of the test c. Select the example of the test d. Design the rubric	

Table 4.2 shows that pre-service teachers had their own strategies in constructing test items. They had around 3-5 strategies. The strategies will

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

be stated in the description with the statement as proof. The first strategy was checking the syllabus or material. This strategy was used by all participants except participant 2 and participant 3. The first strategy used by participant 2 was outlining. Outlining means list kind of questions will be written and kinds of materials will be included in the test. While, participant 3 was making the rubric then checking the material.Based on the result of interview, pre-service teachers argued that checking the material or syllabus can help them to design test based on students' background knowledge. While participant 3 stated that making the rubric in the beginning can help in designing test easily.

"I make rubric first. Then, I choose questions from my references which appropriate with the topic. For example, the rubric is about listening skill so the question is about listening. Then, I choose the topic."(Participant 3)

"Mmm.. I make outline first. So, I list what materials will be included in the test and what kind of questions will be included in the test" (Participant 2)

"First, I look at basic competence or material in that level. Then, I divide the number like one material can be for 5 numbers. For example, the material is recount then I decide how many recount will be in my test" (Participant 4)

"The first thing is checking the material. It means I check what materials have been learned by students. Then I do note taking in the list of material. It makes me easier to decide what materials will be in my test" (Participant 1)

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

The second strategy used by most of pre-service teachers' was looking for some references such as from internet or book. Some of pre-service teachers argued that by looking some references they can get many ideas such as ideas of questions, text, kind of test, and etc. However, participant 2 stated that deciding kind of questions is the second strategy then looking some references is the next strategy. Besides, pre-service teachers argued that these references will be adapted in the test. Some of them preferred to adapt test from internet while others preferred to adapt from book. There were some reasons that make them do adapting. First, some of them argued that it is because the limitation of the time given by lecturer for designing the test. Next, most of them did adapting for the text or audio and participant 1 stated that there will be a new problem if making the text by himself. Similarly, other participants said that it is impossible to create the audio by themselves because they were afraid of making mistakes in the pronunciation.

"I decide kind of questions will be used by me such as fill in the blank, synonym, or antonym. Then I look some references from internet that will be used in my test. I used references which have been implemented in the school because it is trusted. I adapt because I forget the specific lesson in senior high school" (Participant 2)

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

"I adapt some of questions from several books for high school students. Then, I choose which part will be adapted but I usually adapt the text or audio for the questions. It will take a long time if I have to make the text and audio by myself" (Participant 1)

"If get references from my senior high school's book because the level is same. I adapt the context hehe because it makes me easier to design the test(Participant 5)

"I take from others but I modify the questions. I adapt because it is easier hehe. I usually adapt the text but I make the questions by myself"(Participant 4)

The last strategy used by pre-service teachers was starting to design

"After look some references, I design the test by paraphrasing some questions from my references. Then, the last I also make the answer key" (Participant 3)

the test while designing the rubric. After selecting the materials from some references, pre-service teachers started to design the test. Besides, pre-service teachers also re-read test to make sure there is no mistake in the test. After designing the test, they also designed the rubric. However, participant 3 also made the answer key in the last stage.

"I just design the test directly with clear instruction. I must make sure that the instruction is clear." (Participant 4)

"Then, I start writing the questions then the answer depends on the questions. The last, I make the rubric. Actually there must be 'kisi-kisi' for multiple choice but I didn't make it hehe" (Participant 2)

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

> *"After getting examples of test, I select then I start writing the test. The last, I make rubric but it's not totally rubric. For example question 1 until 10 how many marks will be given for these questions."* (Participant 5)

B. Research Discussions

This section describes the main findings of the study and discuss them with relevant theory and previous studies. The discussion focuses directly in the two research questions: challenges and strategies in constructing test items. The detail information of challenges and strategies of test item development is described in following sections.

1. Challenges in Constructing Test Items

The result of this research showed that pre-service teachers encountered different challenges in constructing multiple choice and essay items. The challenges in constructing multiple choice items are writing good sentences for the test, writing the options for multiple choice test, covering all skills in balance, and finding references for the test. Therefore, the challenges in constructing essay items is designing rubric. These challenges of test items construction are explained in the following sections.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

a. Designing Rubric for Open-Ended Questions

In assessing students, pre-service teachers must have rubric to ease in assessing students especially for open-ended questions. Rubric is a range of assessment criteria and an expectation of performance. By having clear criteria, it is expected to assess students objectively. Some of pre-service teachers said that it was challenging when designing the rubric. Mostly, participants argued that the most challenging in designing rubric is writing the criterion for each question objectively and determining the score for each question. Besides, it happened because participants did not have experience before in designing the rubric. Some challenges may be faced in designing rubric. For example, specifying criteria and standards clearly. Second, it includes different dimensions of learning outcomes such as cognitive, psychometric, and affective. The last, rubric is inherently indeterminate when translating judgment on each criterion[50]. Moreover, there was a research which analyzed teachers' competence on developing rubric. The result showed that teachers' rubric was good because it had included almost all criteria.

[50] https://teaching.unsw.edu.au/assessment-rubrics

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

However, teachers still need to improve on some points such as criteria, description, and other components of the rubric[51].

b. Writing good sentences for the test

In writing good sentences for the test, pre-service teachers should pay attention on the grammar and the structure of the sentence. It was hard enough for pre-service teachers because they have not experience yet in doing it. Although, pre-service teachers understand well about grammar but they hesitated that students will understand the meaning of the sentence. However, it is important to write test using good sentence because the sentence of the test can influence students' respond in doing test. According to Board of Scientific Affairs of the American Psychological Association, anxiety, motivation, and performance are affected by the sentences of test items and format[52]. Besides, as stated by Risalatil Umami grammar is the most difficult thing in constructing reading test items[53]. It indicates that writing test for students must consider about grammar and the structure of the test.

---

[51] Silvy Millata, Thesis: *"An Analysis of The Teacher Competence On Developing Rubric To Assess English Competence of Eighth-Grade Students In SMPN 1 Sedati Sidoarjo (Case Study)".* (Surabaya: UINSA 2016)
[52] Board of Scientific Affairs of the American Psychological Association
[53]Risalatil Umami, Thesis: "*Students Ability in Constructing Reading Question Items in Critical Reading Class".* (Surabaya: UINSA 2016)

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

c. Writing the options for multiple choice test

The finding indicated that some of pre-service teachers were aware that writing the choice for multiple choice test was not easy. Most of participants were confused to determine the choice because there are several things to be considered.

1) The stem should be meaningful. It means the stem should present a clear problem. It will be easier to focus on the learning outcome. However, a stem which does not present a definite problem may affect on students' learning outcome.

Which of the following is a true statement ?
A. Procedure text is a text which aims to explain something happened in the past.
B. Procedure text is a text which aims to explain how to make something.
C. Procedure text uses simple past

Figure 4. 1 example of stem which is not meaningful

What are characteristic of procedure text?
A. It has three generic structures such as goals, materials, and steps.
B. It explains about something in the past
C. It uses past continuous

Figure 4. 2 example of stem which is meaningful

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

2) The stem should be contain relevant material. A stem which is not contain relevant material can decrease the validity of the score and the reliability.

> Procedure text is a text that aims to explain about how to make something. This text can be a tips to do something. There are some characteristics of procedure text. For example, using verb in the sentence.**What is the generic structure of procedure text?**
> A. Orientation
> B. Materials
> C. Steps

Figure 4. 3 example of stem irrelevant material

3) The stem should be negatively stated only when significant learning outcomes require it. It is suggested to not using negative phrase because students often get difficulty in understanding the items.

> Which of the following is **not true** about narrative text?
> A. It uses adjective
> B. It tells about imagination
> C. It uses simple present

Figure 4. 4 example of negative phrasing

> Narrative text is a text which is not in a real life. It uses adjective to describe something. There are some generic structures of narrative text, **except**
> D. Orientation
> E. Evaluation
> F. Event

Figure 4. 5 example of negative phrasing better

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

4) The stem should be a partial sentence or question. A question is more acceptable because it will allow students to directly focus on answering the questions.

In addition, _____ is one of generic structure of recount text?

A. Steps
B. Orientation
C. Complication

Figure 4. 6 example of interior blank

In addition, which one is the generic structure of recount text?

A. Steps
B. Orientation
C. Complication

Figure 4. 7 example of partial question

5) No complex multiple choice items. Sometimes, testers use complex options. It makes students confused to answer the questions.

Who is the text about?

A. An talented animator
B. Disney Animator
C. An animation developer
D. A and C
E. B and A

Figure 4. 8 example of multiple choice items

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

The finding shows that most of participants got difficulty in avoiding complex choice, writing stem which is meaningful, and choosing the words for the choice[54].

d. Covering all skills in balance

Skill is one aspect must be considered before constructing the test. Pre-service teachers may decide to include or exclude some skills in English. The finding showed that most of pre-service teachers got difficulties in this area because lack of references. Mostly, the references only provide reading and writing skills. As stated by Harmer J, the elements of the test must be balanced. Elements refers to how many numbers of multiple choice test and essay will be included in the test. Besides, skills included in the test must be balance. It is to make sure that all skills have been mastered by students.

2. Pre-Service Teachers' Strategies In Constructing Test Items

Based on the finding, the strategies used in constructing multiple choice items and essay are same. Some strategies are used by pre-service teachers in constructing test items. The information was gotten through interview the

[54]https://cft.vanderbilt.edu/guides-sub-pages/writing-good-multiple-choice-test-questions/

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

participants. From the finding, the strategies are correlated with theories and it will be explained in below.

a. Check the material

This strategy will help pre-service teachers to determine what materials or topic will be included in the test. Besides, it helps pre-service teacher to choose topic which appropriate for students. It also helps pre-service teachers to design test based on what students have been learned or students' background knowledge. For example, before designing test, pre-service teachers have to look at the syllabus to make sure that the materials are appropriate for students. Based on the finding, pre-service teachers check the syllabus before designing the test. It eased them to design the test. Besides, pre-service teachers also consider the skill that students' have been learned. So, they can decide to include or exclude skills such as speaking. According to Harmer, understanding the syllabus is important. The use of syllabus is to check students' material[55].

[55] Harmer J, *The practice of English language teaching* (Harlow: Pearson Education Limited,2004), 386

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

b. Make an outline

After checking the materials, pre-service teachers also made an outline before designing the test. Harmer mentioned that before designing the test, pre-service teachers must have a list of materials and list of skills will be tested to the students[56]. In addition, Brown stated that the importance of outline is to have a list of skills to be included and how the items will look like[57]. Furthermore, pre-service teachers also determine how many numbers for each material and each skill. In making the outline, pre-service teachers also used internet and book for the references. Pre-service teacher took some questions, text, pictures, and etc to be adapted in the test. As stated by Steven J, test must be legal and proper questions but it does not mean that all test items must be originally constructed[58]. Pre-service teachers made an outline by note taking. Note taking is very useful to make easier in making the outline. Harmer also stated that to construct test items, pre-service teachers also

[56] Harmer J, *The practice of English language teaching* (Harlow: Pearson Education Limited,2004), 386
[57] H Douglas Brown, *Language Assessment: Principle and Classroom Practices*. (USA: Longman. 2004), 50
[58]Steven J Osterlind, *Constructing Test Items: Multiple-choice, Constructed-Response, Performance, and Other Formats* (New York: Kluwer Academic Publishers,2002), 40

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

have to consider time will be given to students while doing test and where students will take the test.

c. Write the test

One criterion of good test items is tests are well written[59]. In writing the test pre-service teachers consider several things such as grammar, punctuation, vocabulary, and etc[60]. It is important to be considered because recent research had revealed that the sentences, the instruction and the format of test items can influence students' respond in doing the test[61].It can be misunderstanding if the instruction and the sentence is not clear enough. Based on the finding, most of pre-service teachers did paraphrase from some references in designing the test. However, pre-service teachers must select which references can be adapted in the test.

d. Make the rubric

In giving the score, pre-service teachersconsider about the level of difficulty in each questions. For instance, the score given for multiple choice questions must be different with essay. Therefore, pre-service

[59]Steven J Osterlind, *Constructing Test Items: Multiple-choice, Constructed-Response, Performance, and Other Formats* (New York: Kluwer Academic Publishers,2002), 40
[60]Steven J Osterlind, *Constructing Test Items: Multiple-choice, Constructed-Response, Performance, and Other Formats* (New York: Kluwer Academic Publishers,2002), 26-30
[61] Board of Scientific Affairs of the American Psychological Association

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

teachers made rubric to ease in assessing students. The benefit of making rubric is to help pre-service teachers grade students' work clearly and efficiently. Besides, students can understand where to focus in the learning efforts. Brown stated that scoring is also important thing in constructing test items. The perception of students' success depend on how many marks is given in each section of the test[62].

[62] H Douglas Brown, *Language Assessment: Principle and Classroom Practices*. (USA: Longman. 2004), 411

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

CHAPTER V

CONCLUSION AND SUGGESTION

In this chapter, the conclusion and the suggestion of this research will be presented below.

A. Research Conclusion

There are two main points that can be concluded from this research as follows:

1. Pre-service teachers' challenges in constructing test items

Pre-service teachers' challenges in constructing multiple choice items are (1) writing good sentences for the test (2) writing the options for multiple choice test (3) covering all skills in balance (4) finding references for the test. Therefore, there is only one challenge in constructing essay items such as designing rubric. Those challenges are divided into 3 aspects such as Knowledge and Experience, Language, and Source. Knowledge and experience refers to pre-service teachers' background knowledge in assessment area such as participants' experience in designing test. Language defined as participants' ability in English skills such as writing, reading, and others. It is important to be mastered by participants especially writing skill because participants must write the sentences of test clearly.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Source refers to participants' ability in selecting appropriate materials for students. The most common challenges occurred by pre-service teachers are designing rubric for open-ended questions. The second aspect mostly faced by pre-service teachers is writing the options for multiple choice test. The last aspect is covering all skills in balance.

2. Pre-service teachers' strategies in constructing test items

Pre-service teachers' has some different strategies in constructing test items. However, pre-service teachers use four strategies. The first is checking the material. Pre-service teachers check students' material before writing the test. It is important to make sure that materials are appropriate for students. The second is making an outline. Outline is very useful for constructing test items because it can ease pre-service teachers in the process of writing test. The next is writing the test clearly. In writing the test, pre-service teachers must consider about the grammar, punctuation, and spelling. Those aspects are important because it can affect students' responses in doing the test. The last is making rubric to make easier in giving students' marks. Making rubric is also an important part in constructing test. It is used to give students' marks for each questions.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

B. Research Suggestion

According on the result of the study, the researcher found some suggestions which follows:

1. For pre-service teachers

   This research informs pre-service teachers who take Language and Assessment Course about the strategies in constructing test items. It includes strategies on how to design rubric and design multiple choice test. Furthermore, this research also provides some challenges mostly faced by pre-service teachers in constructing test items. Pre-service teacherscan learn from the challenges. So, they will not get the same problems as revealed in this research.

2. For Lecturer

   For lecturer of Language and Assessment Course, the lecturer may explain detail information related to challenges faced by pre-service teachers. For example on how to design rubric for open-ended questions and how to write the options for multiple choice test.

3. For the future researchers

   The future researcher may analyze the things that does not mention in this research. The researcher may analyze the test made by pre-service teachers

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

based on the criteria of good test. The similar topic can be conducted for the next research about assessment for more specific.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

REFERENCES

Biggs, J. 1998, *Assessment and classroom learning: A role for summative assessment?*. Assessment in Education: Principles, Policy & Practice, Vol. 5, No. 1, pp. 103-110.

Brown, H. D. 2004. *Language Assessment: Principle and Classroom Practices*. USA: Longman.

Brown, H. D.2001. *Teaching by Principles: An Interactive Approach to Language Pedagogy*. New York: Addison Wesley Longman.

Dr.K.Anandan. *Assessment for learning.* Bharatidasan University, 2015.

Fadhila, Nur Amalia., Thesis: *"Student Teachers' Ability in Designing Assessment Instrument at English Teacher Education Department UIN Sunan Ampel".* Surabaya: UINSA, 2015.

*For Common Core Assessments.*Pearson, 2010.

Harmer J.*The practice of English language teaching*. Harlow: Pearson Education Limited,2004.

Hughes, Arthur. *2007. for Language Teacher*. New York: Cambridge University Press.

J.R. Raco. *Metode Penelitian Kualitatif*. Jakarta: Gramedia Widiasarana Indonesia,2010.

Jabbarifar, Taghi., "*The Importance of Classroom Assessment and Evaluation in Educational System*". Proceedings of the 2nd International Conference of Teaching and Learning, 2009.

John W Creswell. *Educational Research Planning.*UK Sage Publication, 2012.

Michael Bay Borelli.*Considerations For Developing Test Specifications*

Mintzberg Henry. *The rise and Fall of Strategic Planning,* 1998.

Mun.Fika et al. *Complete English Grammar.*Surabaya: Apollo, 1991

Novita, Dian., Thesis: *"An Analysis of Reading Comprehension Questions Made By English Teacher In SMAN 2 Sidoarjo Based on Barret's Taxonomy".* Surabaya: UINSA, 2016.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Osterlind Steven J. *Constructing Test Items: Multiple-choice, Constructed-Response, Performance, and Other Formats*. New York: Kluwer Academic Publishers,2002.

Rahman, Arinal., Thesis: "*An analysis of test items quality in English final examination at Language Development Center".* Banjarmasin: IAIN Antasari 2016

Rea-Dickens and Germanie, 1993; Genesee and Upshur, 1996; O'Mally and Valdez-Pierce, 1996.

Sartika, Nur., Thesis: *"An Analysis of Test Items Based on the Criteria of Good Tests. Bandar Lampung: Lampung University, 2015.*

Shaumu, Yumnaa., Thesis: "*An Analysis of Multiple Choice Items of Computer Based English Test at Smp Alkhairiyah Surabaya*".Surabaya: UINSA 2019.

Suharsini Arikunto. *Prosedur Penelitian Suatu Pendekatan Praktek.* Jakarta : Rineka Cipta,2002.

Toynbee, J. Arnold. *A Study of History.* London: Oxford University Press, 1987.

Umami, Risalatil., Thesis: "*Students Ability in Constructing Reading Question Items in Critical Reading Class".*Surabaya: UINSA, 2016.

Uwe Flick. *An Introduction to Qualitative Research*, 4th ed. Los Angeles: Sage Publications, 2009.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id