AN ANALYSIS OF MULTIPLE CHOICE ITEMS OF COMPUTER BASED ENGLISH TEST AT SMP ALKHAIRIYAH SURABAYA

THESIS

Submitted in partial fulfillment of the requirement for the

degree of Sarjana Pendidikan (S.Pd) in Teaching English

By
Yumnaa Shaumu Istiqoomah
NIM D95214120

ENGLISH TEACHER EDUCATION DEPARTMENT
FACULTY OF TARBIYAH AND TEACHER TRAINING
UNIVERSITAS ISLAM NEGERI SUNAN AMPEL
SURABAYA
2019

PERNYATAAN KEASLIAN TULISAN

Saya yang bertanda tangan dibawah ini:

Nama : YUMNAA SHAUMU ISTIQOOMAH

Nim : D95214120

Jurusan/Program Studi : PENDIDIKAN BAHASA INGGRIS

Fakultas : TARBIYAH DAN KEGURUAN

Dengan ini menyatakan dengan sebenarnya bahwa skripsi yang berjudul *"An Analysis of Multiple Choice Items of Computer Based English Test at SMP AlKhairiyah Surabaya."* Adalah benar-benar merupakan hasil karya sendiri. Segala materi yang diambil dari karya orang lain hanya digunakan sebagai acuan dengan mengikuti tata cara dan etika penulisan karya ilmiah yang ditetapkan oleh jurusan.

Demikian pernyataan ini dibuat dengan sebenar-benarnya. apabila pernyataan tidak sesuai dengan fakta yang ada, maka saya selaku penulis bersedia dimintai pertanggungjawaban sesuai ketentuan peraturan perundang-undangan yang berlaku.

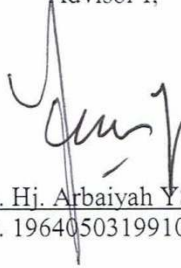Surabaya, 24 Januari 2019

Pembuat Pernyataan,

Yumnaa Shaumu Istiqoomah
NIM.D95214120

ii

ADVISOR APPROVAL SHEET

This thesis by Yumnaa Shaumu Istiqoomah entitled *"An Analysis of Multiple Choice Items of Computer Based English Test at SMP AlKhairiyah Surabaya"* has been approved by the thesis advisors for further approval by the Board of Examiners.

Surabaya, January 24th, 2019

Advisor I,

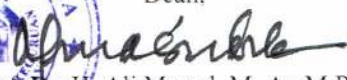Dra. Hj. Arbaiyah YS, MA
NIP. 196405031991032002

Advisor II,

Fitriah, Ph. D
NIP. 197610042009122001

EXAMINER APPROVAL SHEET

This thesis by Yumnaa Shaumu Istiqoomah entitled "*An Analysis of Multiple Choice Items of Computer Based English Test at SMP AlKhairiyah Surabaya*" has been examined on January 24th, 2019 and approved by the Board of Examiners.

Dean,

(Prof. Dr. H. Ali Masud, M. Ag.M.Pd.I)
NIP. 196301231993031002

Examiner I,

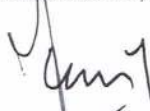(Dr. Mohamad Salik, M.Ag.)
NIP. 196712121994031002

Examiner II,

(Drs. Muhtarom, M.Ed, Gred, Dip Tesol)
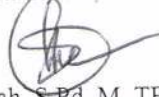NIP. 196512201992031005

Examiner III,

(Dra. Hj. Arba'iyah YS., MA)
NIP. 196405031991032002

Examiner IV,

(Dr. Siti Asmiyah, S.Pd, M. TESOL)
NIP. 197704142006042003

iv

LEMBAR PERNYATAAN PERSETUJUAN PUBLIKASI
KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS

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

Nama : Yumnaa Shaumu Istiqoomah
NIM : D95219120
Fakultas/Jurusan : Tarbiyah / Pendidikan Bahasa Inggris
E-mail address : yumnaistiqomah@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 :

AN ANALYSIS OF MULTIPLE CHOICE ITEMS
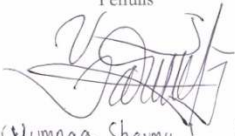OF COMPUTER BASED ENGLISH TEST AT
SMP ALKHAIRIYAH SURABAYA

beserta perangkat yang diperlukan (bila ada). Dengan Hak Bebas Royalti Non-Eksklusif 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, 11 Februari 2019

Penulis

(Yumnaa Shaumu)
nama terang dan tanda tangan

v

ABSTRACT

Istiqoomah, Yumnaa Shaumu. (2019). *An Analysis of Multiple Choice Items of Computer Based English Test at SMP AlKhairiyah Surabaya.* A Thesis. English Teacher Education Department, Faculty of Education and Teacher Training, Universitas Islam Negeri Sunan Ampel Surabaya. Advisor I: Dra. Arbaiyah Yusuf, M.A. Advisor II: Fitriah, Ph. D

Key Words: Computer Based English Year-End Test, Multiple Choice Format-Construction, Multiple Choice Stem-Construction, Multiple Choice Options-Construction, Criteria Validated Multiple Choice Items-Construction

Computer Based English Year-End Test is a part of achievement test that used to measure students' English language achievement at the end of the second semester. This research analyzed the use of the multiple choice items of Computer Based English Year-End Test that was designed by English teacher at SMP AlKhairiyah Surabaya. The research analyzed the constructions of multiple choice items, particularly the multiple choice construction includes 3 categories, those are: format concerns, the writing of the stem, and the writing of the options. In this case, the multiple choice items were expected to have a good construction, since it is designed by the English teacher of school which prioritized the mastery of students in English language. This research focused on analysing the multiple choice format, stem, and options by using criteria of validated multiple choice items construction. The researcher determined this research as content analysis research using descriptive qualitative approach. There is 1 document which became the subject of this research, there was Computer Based English Year-End Test along with its construction. In addition, documentation was used as the techniques to achieve the aims of this research. The analysis result showed that from all of 50 multiple choice items there is none of the items fulfills all the criteria of multiple choice format construction, there is only 1 item fulfills all the criteria of multiple choice stem construction, and there is none of the items fulfills all the criteria of multiple choice option construction. Therefore, it can be concluded that none of the 50 multiple choice items has good multiple choice format and the writing of the options construction.
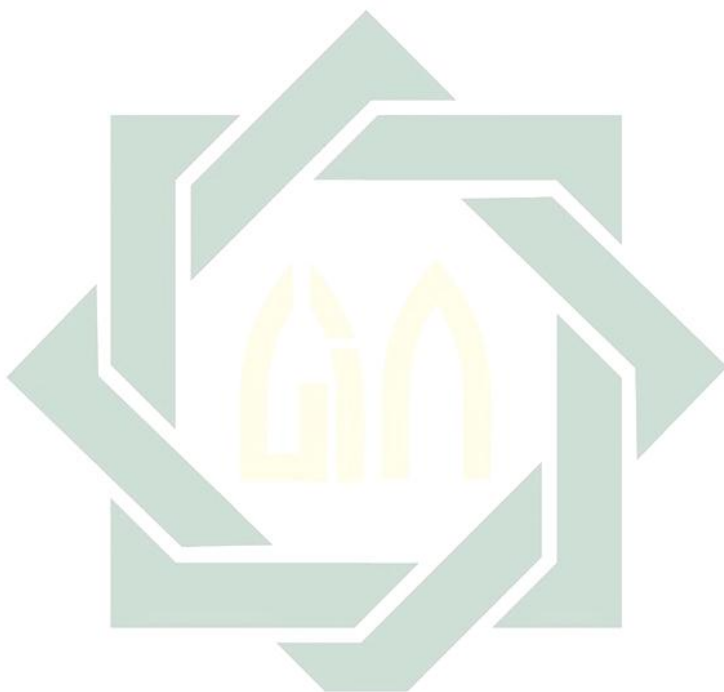
ABSTRAK

Istiqoomah, Yumnaa Shaumu. (2019). *An Analysis of Multiple Choice Items of Computer Based English Test at SMP AlKhairiyah Surabaya.* Skripsi. Pendidikan Bahasa Inggris, Fakultas Tarbiyah dan Keguruan, Universitas Islam Negeri Sunan Ampel Surabaya. Pembimbing I: Dra. Arbaiyah Yusuf, M.A. Pembimbing II: Fitriah, Ph. D

Kata kunci: Penilaian Akhir Tahun Bahasa Inggris Berbasis Komputer, Format Soal Pilihan Ganda, Pokok Soal Pilihan Ganda, Pilihan Jawaban Soal Pilihan Ganda, Kriteria Konstruksi Soal Pilihan Ganda yang Tervalidasi

Penilaian Akhir Tahun bahasa Inggris berbasis Komputer adalah bagian dari tes prestasi yang digunakan untuk mengukur prestasi siswa pada akhir semester kedua. Penelitian ini menganalisis penggunaan soal pilihan ganda pada tes Penilaian Akhir Tahun bahasa Inggris berbasis Komputer yang dirancang oleh guru bahasa Inggris di SMP AlKhairiyah Surabaya. Penelitian ini menganalisis konstruksi soal pilihan ganda, khususnya konstruksi soal pilihan ganda yang meliputi 3 kategori, yaitu: format soal, penulisan pokok soal, dan penulisan pilihan jawaban soal. Dalam hal ini, item pilihan ganda diharapkan memiliki konstruksi yang baik, karena dirancang oleh guru bahasa Inggris di sekolah yang memprioritaskan penguasaan siswa dalam ber-bahasa Inggris. Penelitian ini berfokus pada menganalisis format soal pilihan ganda, pokok soal, dan pilihan jawaban soal dengan menggunakan kriteria konstruksi soal pilihan ganda yang tervalidasi. Peneliti menentukan penelitian ini sebagai penelitian analisis konten menggunakan pendekatan deskriptif kualitatif. Ada 1 dokumen yang menjadi subjek penelitian ini, yaitu Penilaian Akhir Tahun Bahasa Inggris berbasis Komputer beserta konstruksinya. Selain itu, dokumentasi digunakan sebagai teknik untuk mencapai tujuan penelitian ini. Hasil analisis menunjukkan bahwa dari semua 50 item pilihan ganda tidak ada satupun soal yang memenuhi semua kriteria konstruksi format soal pilihan ganda, hanya ada 1 soal yang memenuhi semua kriteria konstruksi pokok soal pilihan ganda, dan juga tidak ada satupun soal yang memenuhi semua kriteria konstruksi pilihan jawaban soal pilihan ganda. Oleh karena itu, dapat disimpulkan bahwa tidak ada satupun dari 50 soal pilihan ganda yang memiliki

konstruksi format dan penulisan pilihan jawaban soal pilihan ganda yang baik, namun hanya 1 soal yang sudah memiliki konstruksi pokok soal pilihan ganda yang baik.

TABLE OF CONTENTS

x

LIST OF TABLES

xi

LIST OF FIGURES

xiii

xiv

xv

xvii

LIST OF ABBREVIATIONS

CBT   Computer Based Test

PBT   Paper Based Test

MCQs   Multiple Choice Questions

xviii
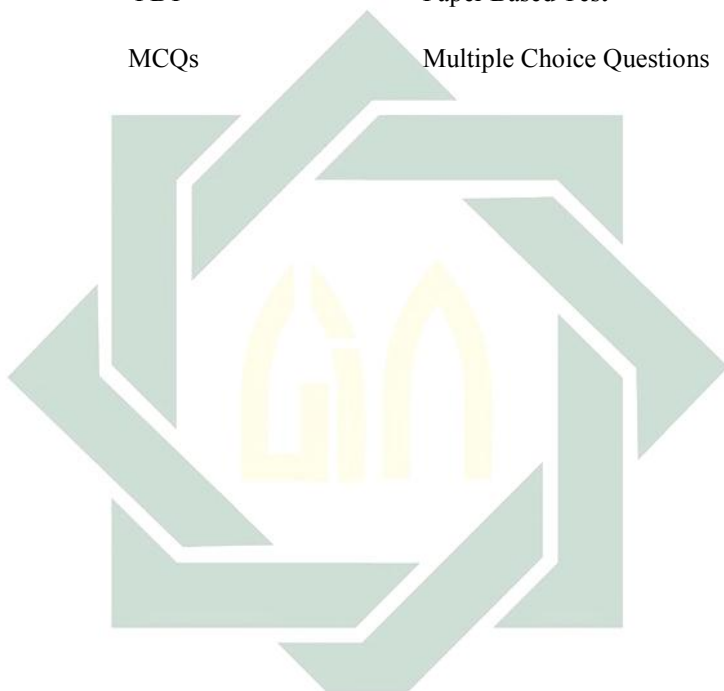
CHAPTER I
INTRODUCTION

This chapter presents the background of the research that describes the reasons why the researcher intends to carry out an analysis of multiple choice items-construction of Computer Based English Test. Then, the problem is formulated in the research question along with the objective of the research. Continue with the significances of the research to inform the usefulness of this research. The limits of what is focused on in this research are written in the scope and limitation of the research. The last part is the definition of key terms used in this research.

A. Background of The Research

The process of teaching and learning in school is not only a mechanism, but also has visions, missions, and goals as targets to be achieved.[1] In achieving these targets, it is necessary to assess the learning as a benchmark to know to what extent the level of competency that the students achieved from the material that have been studied. Collecting information about students' learning outcomes requires methods and instruments, as well as analytical procedures, it is called assessment.[2]

The assessment is an important part of the series of learning processes in education.[3] Therefore assessment of learning outcomes can also determine whether education is good or not. Thus, the accuracy of the assessment of learning outcome has very significant impact on efforts to improve the quality of education in school.[4] Due to the accuracy of the assessment of learning outcome has a very significant impact on effort to improve the quality of education in school. So, the quality of language test as one of the assessment tool really needs to be considered.

The language test is divided into several forms, one of which is multiple-choice item.[5] Multiple choice items are often created and

---

[1] H. Douglas. Brown, *Language Assessment Principles and Classroom Practice* (San Fransisco : Longman, 2004), 4.

[2] Direktorat Pembinaan SMP Tahun 2017, Kemendikbud, *Penilaian oleh Pendidik dan Satuan Pendidikan Sekolah Menengan Pertama* (Jakarta : Kemendikbud, 2017), 7.

[3] M. David Miller - Robert L. Linn, *Measurement and Assessment in Teaching, Tenth-Edition* (New Jersey : Pearson Education, Inc.,Upper Saddle River, 2009), 26.

[4] Claire Wyatt, et.al., *Educational Assessment in the 21st Century: Connecting Theory and Practice* (London : Springer, 2009), 45.

[5] Claire Wyatt, et.al., *Educational Assessment in the 21st Century ....,* 55.

used by teacher to assess student learning outcomes. Since multiple choice items can be used to test a greater variety of instructional objectives.[6] The multiple-choice item is also one of the most popular item formats used in educational testing. Many group-administered standardized tests consist of multiple-choice item.[7]

The teacher has an important role to measure the success of the students in achieving the learning target after the learning process. It is because the teacher determines the learning objectives, formulates the basic competencies and indicators to be taught, selects the teaching materials, conducts the learning activities, and assesses the learning outcomes of the students, so that the teacher knows the most about what competencies should be measured from the students.[8] Based on this situation, English teacher in SMP AlKhairiyah Surabaya make their own multiple choice item which is used as a tool to assess students' learning outcomes in assessment.[9] This makes the teacher's made multiple choice test plays an important role in the process of retrieving information about the achievement of students' competency.

One example of multiple choice items at SMP AlKhairiyah Surabaya is Year-End Test (PAT).[10] This is an activity undertaken by educator at the end of second semester.[11]. Year-End Test (PAT) is an important educational assessment tool.[12] The purpose of Year-End Test (PAT) is measuring the achievement of students' learning outcomes that is the goal of instruction.[13]

Year-End Test (PAT) of SMP AlKhairiyah Surabaya uses Computer Based Test (CBT). In this modern era, it is a familiar

---

[6] Steven J. Osterlind, *Constructing Test Items: Multiple-Choice, Constructed-Response, Performance, and Other Formats, Second-Edition* (Moscow : Kluwer Academic Publisher, 2002), 163.

[7] Albert Oosterhof, *Developing and Using Classroom Assessment, Third-Edition* (New Jersey : Merrill/Prentice Hall, 2003), 85.

[8] William. A. Mehrens - Irvin J. Lehmann, *Measurement and Evaluation in Education and Psychology, Fourth-Edition* (Belmont : Wadsworth, Thomson Learning, 1991), 51.

[9] *The Result of Interview with English Teacher of Grade IX - Mrs. Nyimas, on Wednesday, 30th July 2018, at 14.00 – 15.30 WIB* (Surabaya : SMP AlKhairiyah Surabaya, 2018)

[10] *The Result of Interview with English Teacher of Grade IX - Mrs. Nyimas, ...*

[11] Permendikbud, Kemendikbud, *Standar Penilaian Pendidikan* - BAB V Pasal 6: Bentuk Penilaian (Jakarta: Kemendikbud, 2016), 5.

[12] Direktorat Pembinaan SMP Tahun 2017, Kemendikbud, *Penilaian ...*, 71.

[13] Thomas. M. Haladyna, *Developing and Validating Multiple-Choice Test Items, Third-Edition* (New Jersey : Lawrence Erlbaum Associates, 2004), 6.

thing to use a computer as a medium for assessing. Computer Based Test (CBT) is a modern test that uses the application on computer as its medium.[14] In responding to this positive thing in the field of education, Indonesia applying computer-based test method.[15] This is due to obtaining accuracy and reliability of marking calculation.[16] The use of Computer Based Test (CBT) in Indonesian assessment is already conducted in many school, one of them is SMP AlKhairiyah Surabaya. Here has already used it on various types of assessment. So, students are expected to be easier in using Computer Based Test (CBT) especially for facing Computer Based National Examination (UNBK) later.

Based on the preliminary research with an English teacher at SMP AlKhairiyah Surabaya, it is known that English Year-End Test (PAT) score of some students at grade VIII in the 2017/2018 academic years is under the Minimum Exhaustiveness Criterion (KKM), that is 77. There are some students who follow remedial test because the score is still below the standard. Students' level achievement (KKM) known by seeing the shown score after doing the test on the computer automatically.[17]

The problems that occur at SMP AlKhairiyah Surabaya can not be underestimated, if the multiple choice items used to measure the achievement of the students' competence have not been analyzed and developed before. So, it may affect the quality of the item itself. The test uses Computer Based Test (CBT), it may issue some problem than Paper Based Test (PBT). The most primary issue here is whether computer administration of a multiple choice item produces the same result as Paper Based Test (PBT).[18] Therefore, multiple choice items should be analyzed before it is used.[19]

English teacher at SMP AlKhairiyah Surabaya has never done the item analysis, the teacher only made multiple choice items for Year-End Test in accordance with all basic competence materials taught during second semester. The multiple choice test made by

---

[14] H. Douglas. Brown, *Language Assessment Principles and Classroom Practice* ...., 14.
[15] Direktorat Pembinaan SMP Tahun 2017, Kemendikbud, *Penilaian* ...., 101.
[16] Direktorat Pembinaan SMP Tahun 2017, Kemendikbud, *Penilaian* ...., 26.
[17] *The Result of Interview with English Teacher of Grade IX - Mrs. Nyimas*, ...
[18] Thomas. M. Haladyna, *Developing and Validating Multiple-Choice Test* ...., 93.
[19] J.B. Heaton, *Writing English Language Tests* (New York : Longman, 2000), 7.

the teacher here is constructed without using the guidelines of how to construct multiple choice items.[20] It means that the quality of multiple choice construction has not been known whether it includes the multiple choice items that qualify as a good test which is written using validated multiple choice item-construction guidelines or not.[21] Since, this situation can have an impact on misinterpretation of student achievement outcomes. In other word, it will have an effect of error tendency in interpretation of test results, because the error construction of the test items may make the students more difficult to response (answer) the item, not perform according to the expectations.[22]

This error may come from three components of multiple choice items those are format concerns, the writing of stem, and the writing of the options.[23] The format concerns is related to the layout or the arrangement of the writing, the writing of the stem is related to how to write a part of multiple choice item (the question of the test), and the writing of options is related to how to write a part of multiple choice item (the answer-choices) to response the question. These components help the teacher in designing the contruction of multiple choice items. Despite, the error of these components certainly has an impact on the biased information of true students ability obtained from the test.

The absence of attempts to analyze this test item can be due to the several factors, such as the teacher's tendency to ignore the importance of knowing the quality of teacher-made test construction, the lack of evaluation planning, and the lack of teacher mastery towards how to analyze the test item construction.[24] By finding the error construction, it may become the references to be aware of the previous fault in constructing multiple choice items. Therefore, item analysis should be made in preparing teachers' multiple choice items in a structured and planned construction in order to obtain an objective and accurate assessment result. There

---

[20] *The Result of Interview with English Teacher of Grade IX - Mrs. Nyimas*, ...
[21] Steven M. Downing - Michael C. Rodriguez, *A Review of Multiple Choice Item-Writing Guidelines for Classroom Assessment* (New Jersey : Lawrence Erlbaum Associates, 2004), 326.
[22] J.B. Heaton, *Writing English Language Tests* ...., 37.
[23] Thomas. M. Haladyna, *Developing and Validating Multiple-Choice Test* ...., 97.
[24] *The Result of Interview with English Teacher of Grade IX - Mrs. Nyimas*, ...

are many item analysis that have been made to help in analyzing items and have been done as much as the results of the previous studies. These methods are very helpful for the teacher in analyzing the items easily.

The first study comes from David DiBattista, under the title "*Examination of the Quality of Multiple-choice Items on Classroom Test*".[25] This study aims to determine the characteristics of the items include the discrimination index, and the effectiveness of distractor. This study used statistical item analysis. The research data was analyzed statistically using the help of undergraduates' examinees responses to 1198 multiple-choice items. This study show that the quality of many multiple choice items are still far from the perfection and suggested to be improved by conducting the item analysis and by modifying distractors that impair the discrimination index of items.

The second study comes from Dian Nurliyanto, under the title *'Test Item Analysis of The Final Examination on Economics at Grade XII SMA Negeri Banyumas in the 2014/2015 Academic Years'.*[26] This study uses Analysis of items including difficulty index, discrimination index and pattern of answer distributions according to Anas Sudjono. The purpose of this study was to determine the quality of the first semester of Final Examination on Economics at Grade XII SMA Negeri Banyumas in the 2014/2015 academic years in terms of difficulty index, discrimination index, and the pattern of answers distribution. This study shows that test items are prepared as simple as possible without using good construction criteria and without carrying out an item analysis. These situations made bad quality for test items as well as the result of this study. So, they can be revised and replaced by making new items.

Both of the previous studies show that, if the quality of the items hasnot been known before, it will have an effect of error tendency in interpretation of true students' ability obtained from the

---

[25] David DiBattista, "Examination of the Quality of Multiple-choice Items on Classroom Tests". *The Canadian Journal for the Scholarship of Teaching and Learning.* Vol. 2, Iss. 2, Art. 4, December 2011.
[26] Dian Nurliyanto, Undergraduate Thesis: *"Test Item Analysis of The Final Examination on Economics Subject Grade XII IPS SMA Negeri Banyumas Academic Year 2014/2015"* (Yogyakarta : Yogyakarta State University, 2015)

assessment tool. So, the two previous studies above analyze the test result including the difficulty index, discrimination index, and the effectiveness of distractor. But, they did not analyze the construction of multipe choice items, especially on Computer Based English Test.

This analysis is the most important part of a good test-making procedure. Since, the multiple choice items-construction guidelines using all the strategies possible to make the test item as good as it can be. When multiple choice items are written without regarding validated multiple choice item-construction guidelines, the consequences can be negative.[27] Based on it, the researcher do the analysis of multiple choice items dealing with the format, the writing of stem, and the writing of options by using the criteria of multiple choice items-construction guidelines. This item analysis is to determine whether the construction of multiple choice item fulfilled the criteria of validated multiple choice item-construction guidelines or not to improve the quality and increase the objectivity of it in measuring the students' ability according to expectations. Also by finding the error construction, it may become the references to be aware of the previous fault in constructing multiple choice items.

B. Research Question

In relation to the background of the research above, the problem of this research is "how is the construction of multiple choice items of Computer Based English Test at SMP AlKhairiyah Surabaya?"

C. Objective of The Research

This research was aimed at finding out how the construction of multiple choice items of Computer Based English Test at SMP AlKhairiyah Surabaya is.

D. Significances of The Research

The researcher hopes that the result of this research can be used by the researcher, teacher, students, educational institution, and next researcher for the following purposes:

1. For the researcher, this research is expected to open new insights on the analysis of multiple choice items. This research can be a new experience that researcher can make good

---

[27] Thomas. M. Haladyna, *Developing and Validating Multiple-Choice Test ....*, 97.

construction of multiple choice items to get good quality of multiple choice test.

2. For the teacher, this research can be used as input for the teacher to know the importance of doing multiple choice analysis to make good multiple choice item construction especially for Computer Based English Test at SMP AlKhairiyah Surabaya.

3. For students, this research can give students good quality of test items in accordance with the purpose of learning outcomes.

4. For educational institution, the educational institution can improve the quality of education in the school especially in making good multiple choice items of Computer Based Test at SMP AlKhairiyah Surabaya.

5. For the next researcher, the result of this research can be used as previous research for the next researcher.

E. Scope and Limitation of The Research

1. Scope of The Research

The scope of this research focused on a multiple choice items-construction, which is by analyzing the multiple choice items that have been tested. This research analyze the multiple choice item construction using the criteria of validated multiple choice item-construction guidelines including format concerns of multiple choice items, the writing of stem, and the writing of options by Thomas M. Haladyna.

There are kinds of multiple choice-format to measure knowledge, skills, and abilities.[28] Such as; conventional multiple choice items, matching items, extended matching items, alternate choice items, and true-false items.[29] This research focused on the conventional multiple choice items, since the most common multiple choice format is conventional.[30] Conventional multiple choice item has parts: a stem and options (the correct choice and several wrong answers, called distracters).[31] So, related to the format concerns this research focus on analyzing the lay out or the arrangement

---

[28] Thomas. M. Haladyna, *Developing and Validating Multiple-Choice Test ....*, 41.

[29] Thomas. M. Haladyna, *Developing and Validating Multiple-Choice Test ....*, 67.

[30] Thomas. M. Haladyna, *Developing and Validating Multiple-Choice Test ....*, 67.

[31] Thomas. M. Haladyna, *Developing and Validating Multiple-Choice Test ....*, 68.

of the stem and options of conventional multiple choice item-structure.

The stem should set a complete idea of the knowledge to be indicated by choosing the right answer.[32] It may contain language areas: grammar and vocabulary (concerns on word meaning, word formation, and collocation).[33] In testing those language areas, stem may be written in some forms, those are a incomplete sentence, a complete sentence, and a question. So, related to the writing of stem, this research focus on analyzing the design of language areas and the forms of stem.

The option of multiple choice items consist of one correct answer, and several distractors. The correct option can be a word, phrase, sentence, paragraph and even a drawing or photograph. As well as the distractors are also in those forms. So, related to the writing of options, this research focus on analyzing the forms of options.

Then relate the result of the analysis above to the criteria of multiple choice items-construction guidelines to analyze whether the item-construction fulfilled the criteria of validated multiple choice item-construction guidelines or not, to find how the construction of multiple choice items of Computer Based English Test is.

2. Limitation of The Research

The limitation of this research is to analyze the multiple choice items of Computer Based English Year-End Test at grade VIII SMP AlKhairiyah Surabaya in the 2017/2018 academic years.

F. Definition of Key Terms

In order to have the same idea and concept in this research, the researcher clarifies the terms used in this research, as the details are:

1. Multiple Choice Items

Multiple choice item is an item consists of a stem and several alternatives (option or choices). The stem may be a

---

[32] Thomas. M. Haladyna, *Developing and Validating Multiple-Choice Test ....*, 68.
[33] J.B. Heaton, *Writing English Language Tests ....*, 9.

question or an incomplete statement. The options include the correct answer and several wrong answers called distractors.[34]

2. The Construction of Multiple Choice Items

The construction of multiple choice items is the building or the structure of items that have format aspects, stem and options concern of multiple choice items.[35]

3. Analysis of Multiple Choice Items Construction

An item analysis is the observation that may gauge the quality of items, discard those which are unacceptable, and retain those which meet criteria of merit.[36]

4. Criteria of Multiple Choice Items-Construction

Criteria of multiple choice items-construction guidelines are a set of guidelines should be adopted to construct multiple choice items including format concerns category, the writing of stem category, and the writing of options category. Item-construction guidelines using all the strategies possible to make the test item as good as it can be. When multiple choice items are written without regard for validated multiple choice item-construction guidelines, the consequences can be negative.[37]

5. Computer Based English Year-End Test in SMP AlKhairiyah Surabaya

English Year-End Test is an activity to describe the level of students' achievement of English lesson in the end of second semester. [38] In this research, English Year-End Test (PAT) of SMP AlKhairiyah Surabaya uses Computer Based Test (CBT).[39] Computer Based Test (CBT) is a modern test that uses an application on computer.[40]

---

[34] Norman E. Grondlund, *Assessment of Students Achievement, Seventh-Edition* (Boston : Pearson Education, 2003), 60.

[35] Thomas. M. Haladyna, *Developing and Validating Multiple-Choice Test ....*, 97.

[36] Steven J. Osterlind, *Constructing Test Items: Multiple-Choice ....*, 257.

[37] Thomas. M. Haladyna, *Developing and Validating Multiple-Choice Test ....*, 97.

[38] Permendikbud, Kemendikbud, *Standar Penilaian Pendidikan ....*, 5.

[39] *The Result of Interview with English Teacher of Grade IX - Mrs. Nyimas*, ...

[40] H. Douglas. Brown, *Language Assessment Principles and Classroom Practice ....*, 14.

CHAPTER II
REVIEW OF RELATED LITERATURE

This chapter presents review of related literature which contains the basis theories of assessment and language test, types of test maker, Year-End Test, multiple choice items, Computer Based Multiple Choice Test, good test assembly, criteria of validated multiple choice items-construction, and analysis of multiple choice items construction used in this research background. Continue with review of previous study which describes the differences between this research and the other studies which are done by another researcher.

A. Review of Related Literature

1. Teacher Made Test

Teachers have an obligation to provide their students with the best possible instruction. This implies that they must have some procedures whereby they can reliably and validly evaluate how effectively their students have learned what had been taught.[1] Teachers know better what classroom needs, backgrounds, strengths, and weaknesses of their students.[2] Therefore, the teachers created and developed their own language test that is able to measure the students' ability as accurately as possible, which is called a teacher made test.

This test is also called a non-standard test, because it is usually composed by teachers who do not have sufficient knowledge and experience in the test preparation, or they actually have the skills but are not perfect in preparing, testing, and analyzing the test well. So, the quality of the test items have not be accounted for.[3] The example of teacher made test in Indonesia are Daily Test (PH), Mid-Term Test (PTS), Final Test (PAS), and Year-End Test (PAT), and School Examination (US).[4]

2. Year-End Test

Permendiknas number 23 in 2016 says Year-End Test (PAT) is an educational assessment tool conducted at the end

---

[1] William. A. Mehrens - Irvin J. Lehmann, *Measurement and Evaluation* ...., 50.
[2] William. A. Mehrens - Irvin J. Lehmann, *Measurement and Evaluation* ...., 51.
[3] M. David Miller - Robert L. Linn, *Measurement and Assessment in Teaching,* ...., 401.
[4] Direktorat Pembinaan SMP Tahun 2017, Kemendikbud, *Penilaian* ...., 101.

10

of second semester.[5] This Year-End Test (PAT) has an important role in measuring the achievement of students' learning outcomes that is the goal of instruction or training,[6] also to summarize what students grasped at the end of a unit of instruction.[7]

Scope of Year-End Test item includes all indicators that represent Basic Competencies (KD) in the second semester. Year-End Test (PAT) results, then processed and analyzed to find out the learning completeness of students. The results of this assessment can be used for remedial programs, enrichment and report cards.[8]

3. **Multiple Choice Items**

Tests try to find out about students' knowledge through more controlled items, such as multiple choice items. The multiple choice item is often easier to mark, and produce greater realiability of the score.[9]

Multiple choice item is considered to be ideal test instrument for measuring students' knowledge. The correct answer can be read by computer, it is simply matter of checking the answers for each item, thereby cutting out the possibility of scoring error.[10]

Multiple choice item consists of a stem and several alternatives (options or choices). The stem may be a complete sentence, an incomplete sentence, or a question. The options include the correct answer and several wrong options called distractors.[11] The detail described in the following:

a. **Multiple Choice Format**

Multiple choice items can be called as conventional multiple choice. This item format has parts: stem and options (correct answer; and several wrong answers, called distracters).[12] The figure 2.1 which is written in the

---

[5] Permendikbud, Kemendikbud, *Standar Penilaian Pendidikan ....*, 5.

[6] Thomas. M. Haladyna, *Developing and Validating Multiple-Choice Test ....*, 6.

[7] H. Douglas. Brown, *Language Assessment : Principles and Classroom Practice ....*, 6.

[8] Direktorat Pembinaan SMP Tahun 2017, Kemendikbud, *Penilaian ....*, 100-101.

[9] Jeremy Harmer, *The Practice of Language Teaching, Third-Edition* (London : Longman Group, 2007), 322.

[10] Jeremy Harmer, *The Practice of Language Teaching ....*, 323.

[11] Norman E. Grondlund, *Assessment of Students Achievement ....*, 60.

[12] Thomas. M. Haladyna, *Developing and Validating Multiple-Choice Test ....*, 68.

book *Writing English Language Test*, shows the multiple choice format.[13]

Figure 2.1 Multiple Choice Format

The detail of stem and options described in the following:

1) Stem

The initial part of multiple-choice item called the stem.[14] The stem may be a incomplete sentence, an complete sentences, or a question. The figure 2.2 which is written in the book *Writing English Language Test*, shows same forms of stem.[15]

Figure 2.2 Forms of Stem

---

[13] J.B. Heaton, *Writing English Language Tests* ...., 28.

[14] J.B. Heaton, *Writing English Language Tests* ...., 28.

[15] J.B. Heaton, *Writing English Language Tests* ...., 30.

2) Options

The options from which the students can choose their answers referred to as options.[16]

a) Correct choice

One option as written in the figure 2.2 is the correct answer or key.[17] The correct option is the one and only right answer. The correct answer can be a word, phrase, sentence, paragraph and even a drawing or photograph.[18] In the figure 2.2 with the incomplete, complete and the question stem, the second part of the sentence are the options, and one of these is the right answer.[19] The forms of correct choice written in the figure 2.2 with incomplete sentence is a word, with complete stem is sentence, with a question is sentence.

b) Distractors

The other options are distractors. The task of ditractors is to distract the majority of poor students (those who do not know the answer) from the correct option.[20]

A distractor is an undoubtedly wrong answer.[21] As well as the correct choice, distractors can be a word, phrase, sentence, paragraph and even a drawing or photograph.[22] Figure 2.2 shows multiple choice items with the incomplete, complete and the question stem, the second part of the sentence are the options, and three of these except the correct answer is the distractors.[23]

---

[16] J.B. Heaton, *Writing English Language Tests* ...., 28.
[17] J.B. Heaton, *Writing English Language Tests* ...., 28.
[18] Thomas. M. Haladyna, *Developing and Validating Multiple-Choice Test* ...., 68.
[19] Thomas. M. Haladyna, *Developing and Validating Multiple-Choice Test* ...., 69.
[20] J.B. Heaton, *Writing English Language Tests* ...., 28.
[21] Thomas. M. Haladyna, *Developing and Validating Multiple-Choice Test* ...., 69.
[22] Thomas. M. Haladyna, *Developing and Validating Multiple-Choice Test* ...., 68.
[23] Thomas. M. Haladyna, *Developing and Validating Multiple-Choice Test* ...., 69.

4.   **Computer Based Multiple Choice Test**

In this modern era, it is a familiar thing to use a computer as a medium for assessing. In responding to this matter as a positive thing in the field of education in Indonesia applying computer-based test methods, two of them are National Standard School Examination (USBN) and Computer Based National Examination (UNBK).[24] This is due to obtaining accuracy and reliability of marking calculation.[25]

The computer is currently being used in many areas of assessment. In addition to the already established uses of computers for test scoring, calculation of final grades and test score reporting.[26] Computer-Based Test (CBT) has a number of important advantages compared to traditional Paper-Based Test (PBT) such as efficiency, immediate scoring and feedback in the case of multiple-choice items.[27] The use of Computer Based Test (CBT) is also effective to minimize the usual level of cheating during the test. This can be known from the type of items written randomly. If the Paper Based Test (PBT) can only provide 5 types of items, then for Computer Based Test (CBT) can provide up to 20 types of items, depend on the number of students who take the test.[28] Teachers also do not need to correct the answer sheet because the application of the test items automatically shows the score of students after doing all the items on the computer. This suggests that the use of Computer Based Test (CBT), in addition to being effective is also efficient, both time and cost.[29]

Computer can be used in testing process, but with the use of computer may change the performance characteristics

---

[24] Direktorat Pembinaan SMP Tahun 2017, Kemendikbud, *Penilaian* ...., 101.

[25] J.B. Heaton, *Writing English Language Tests* ...., 26.

[26] Christine Hui Li Chin, *The Effect of Computer Based Tests on the Achievement, anxiety, and attitudes of Grade 10 Sciece Students* (Columbia : The University of British Columbia, 1990), 9.

[27] Anja J. Boeve, *Introducing Computer-Based Testing in High-Stakes Exams in Higher Education: Results of a Field Experiment* (US : US National Library of Medicine National Institutes of Healt, 2015), 1.

[28] *The Result of Interview with English Teacher of Grade IX - Mrs. Nyimas*, ...

[29] *The Result of Interview with English Teacher of Grade IX - Mrs. Nyimas*, ...

of items without computer as Paper Based Test (PBT).[30] The most primary issues here is whether computer administration of a multiple choice item produces the same result as Paper Based Test (PBT).[31] Therefore, multiple choice items construction should be analyzed before the multiple choice item is used.[32]

5. Criteria of Validated Multiple Choice Items-Construction

There are several criteria regarding validated multiple choice construction. A set of criteria should be adopted to construct multiple choice items. Item-construction criteria using all the strategies possible to make the test item as good as it can be. When multiple choice items are written without regard for validated multiple choice item-construction, the consequences can be negative.[33]

Test items need to fulfil the criteria of validated multiple choice item construction in classroom assessments, so that students perform according to expectations. Since, the use of these criteria serve items writers well, particularly teachers who traditionally have difficulty in writing items to assess the student achievement. Therefore assessment activities, especially writing and designing the test tend to have high standards for item development and use these criteria consistently.[34]

These multiple choice item-construction criteria consist of three categories in developing multiple choice items including format concerns-category (designing and formatting the structure of multiple choice items), writing stem category, and writing options category including the right answer and the distractors.[35] The following are the criteria of validated multiple choice items construction:

---

[30] Thomas. M. Haladyna, *Developing and Validating Multiple-Choice Test ....*, 92.
[31] Thomas. M. Haladyna, *Developing and Validating Multiple-Choice Test ....*, 93.
[32] Thomas. M. Haladyna, *Developing and Validating Multiple-Choice Test ....*, 7.
[33] Thomas. M. Haladyna, *Developing and Validating Multiple-Choice Test ....*, 97.
[34] Steven M. Downing - Michael C. Rodriguez, *A Review of Multiple Choice Item ....*, 326.
[35] Thomas. M. Haladyna, *Developing and Validating Multiple-Choice Test ....*, 98.

a. **Format Concerns**

1) Format items vertically instead of horizontally.

The items should be formatted vertically instead of horizontally.[36] Students who may have test anxiety, horizontal formatting may be harder to read, thus confusing students and lowering test scores.

The advantage of horizontal formatting is economy; you can fit more items on a page. If appearance is important, vertical formatting looks less cramped and has a better visual appeal. So, item formatted vertically easier to be read.[37] The book *Developing and Validating Multiple-Choice Test Items (Third-Edition)* presents the example of item formatted horizontally and vertically shown in the figure 2.3.[38]

You draw a card from a deck of 52 cards. What is the chance you will draw a card with an odd number on it?
A.   36/52   B.   32/52   C.   About one half

You draw a card from a deck of 52 cards. What is the chance you will draw a card with an odd number on it?
A.   36/52
B.   32/52
C.   About one half

**Figure 2.3 Item Formatted
Horizontally and Vertically**

2) Edit items for correct punctuation, capitalization, spelling and grammar.

Items should content correct punctuation, capitalization, spelling and grammar.[39] Reordering options with other editorial may affect item characteristics.[40] For example capital letters are only used in options which occur at the beginning of a

---

[36] Thomas. M. Haladyna, *Developing and Validating Multiple-Choice Test ....*, 105.
[37] Thomas. M. Haladyna, *Developing and Validating Multiple-Choice Test ....*, *105.*
[38] Thomas. M. Haladyna, *Developing and Validating Multiple-Choice Test ....*, *105.*
[39] Albert Oosterhof, *Developing and Using Classroom Assessment ....*, 92.
[40] Thomas. M. Haladyna, *Developing and Validating Multiple-Choice Test ....*, 105.

sentence.[41] So, editing the items may be needed. Here are some general rules of punctuation, capitalization, spelling and grammar that can be applied in writing.

a) Punctuation

Some general rules of Punctuation described in the following:

(1) Periods (.)

Some rules of using "Periods" described in the following[42]:

(a) At the end of a declarative sentence (a sentence that makes a statement).

(b) At the end of an imperative sentences (a sentence that makes a request, gives an instruction, or states a command).

(c) At the end of a sentence that asks an indirect question.

(d) After an initial of the name.

(e) After an abbreviation, including titles such as: Mr., Mrs., and Dr.

(f) Before a decimal (also known as a decimal point).

(g) Between dollars and cents.

(h) In place of the dot when writing about the Internet. In other words, when people talk about the Internet, they say "dot," as in "dot-com." When writing about the Internet, the "dot" is a period and is not typically written as a word.

(2) Question Marks (?)

Some rules of using "Question Marks" described in the following[43]:

(a) After an interrogatory sentence (a word or group of words that asks a direct question).

---

[41] J.B. Heaton, *Writing English Language Tests* ...., 32-33.

[42] Brady Smith, *Proofreading, Revising, & Editing Skills* ...., 121-123.

[43] Brady Smith, *Proofreading, Revising, & Editing Skills* ...., 124.

    (b) Sentences that begin with the 5 W's and 1 H of journalism: Who, what, where, when, why and how are usually questions.

(3) Exclamation Points (!)

   Some rules of using "Exclamation Points" described in the following[44]:

 (a) Use an exclamation point after an exclamatory sentence (a sentence that expresses strong feeling).

 (b) Use an exclamation point after an interjection (a word or group of words used to express surprise or other emotion).

(4) Using Commas (,)

   Some rules of using "Commas" described in the following[45]:

 (a) A comma separates an introductory word from the rest of the sentence. Example: Sadly, the summer ended.

 (b) Words such as yes, no, well, why, and oh are also followed by a comma when they begin a sentence.

 (c) Commas with Appositives.

   Appositives are also set off by commas. An appositive is a noun or a pronoun that follows another noun or pronoun to identify or explain it. An appositive phrase is a phrase that contains an appositive. Example: Frances, the best photographer in the state, took pictures at my wedding.

 (d) Items in a Series

   These separated items can be words, phrases, or clauses.

---

[44] Brady Smith, *Proofreading, Revising, & Editing Skills ....*, 125.
[45] Brady Smith, *Proofreading, Revising, & Editing Skills ....*, 128-132.

(e) Dates and Addresses

Put a comma between the date and the year when writing the date in a month-day-year format.

(f) Put commas on either side of the state or province when it is used to further identify a city.

(5) Apostrophes (')

Apostrophes are used to show that things belong to one or more people or things. That is called possession. Example: Shakema's folder.[46]

(6) Quotation Mark (".....")

Some rules of using "Quotation Mark" described in the following[47]:

(a) Direct Quotations

Direct quotations are someone's exact words. Example: The assistant manager said I am a "very quick learner".

(b) Dialogue.

Dialogue is a form of direct quotation because you are writing someone's exact words. Dialogue is a conversation between two or more people.

(c) Titles

Quotation marks are used to punctuate the titles of shorter works, or works that are parts of a larger work. When you refer to a published piece in your writing, be sure to cite your sources. This will eliminate the problem of plagiarism. Example of a chapter title: Your assignment tonight is to read Chapter 12, "Civil Rights in America."

---

[46] Brady Smith, *Proofreading, Revising, & Editing Skills ....*, 143.
[47] Brady Smith, *Proofreading, Revising, & Editing Skills ....*, 147-150.

(7) Single Quotation Marks.[48]

Single quotation marks (') set apart a quotation within a quotation. Example: "Has anybody read 'Do Not Go Gentle into That Good Night' by Dylan Thomas?" asked Ms. Dwyer.

b) Capitalization

Capitalize each word described in the following[49]:

(1) First Word of the Sentence.
(2) Pronoun I.
(3) First word of direct quotation.
(4) Proper Nouns and Proper Adjectives.

Such as: name of a person (family members; ethnic groups, races, languages, and nationalities), thing (brand names of products; days of the week; months; special events and calendar events; holidays, names of trains, ships, and other modes of transportation), and place (names of structures and buildings; cities, states, and governmental units; institutions, organizations, and businesses; public parks and bodies of water; streets, highways, and roads).

c) Spelling and Grammar

To check the error spelling can use two ways, those are Using "Spelling & Grammar" in Ms. Word and Using Dictionary. The following are the explanation about those ways to check the word-spelling:

(1) Using "Spelling & Grammar" in Ms. Word
(a) Checking Spelling

The items-writer may use spell & grammar-checking in Ms. Word. It checks the document-spelling which is

---

[48] Brady Smith, *Proofreading, Revising, & Editing Skills ....*, 151.
[49] Brady Smith, *Proofreading, Revising, & Editing Skills ....*, 109-110.

typed defaultly. Ms. Word uses a red squiggly line to underline the words which has wrong spelling. Ms. Word makes some corrections immeadiately. These corrections can be called as AutoCorrect.[50]

AutoCorrect fixes some of typos or wrong spelling. For example, if you type "spelling", Ms. Word automatically changes it into "spelling". It can easily check spelling of words that are not included in Ms. Word's dictionary. To check a single word, click the word and press "F7" to open the Spelling and Grammar dialog box. The figure 2.4 which is written in the book *Absolute Beginner's Guide to Microsoft Office*, shows the spelling and grammar checking dialog box.[51]

Figure 2.4 Spelling and Grammar Checking Dialog Box

---

[50] Jim Boyce, *Absolute Beginner's Guide to Microsoft Office* (Indiana: Que Publishing, 2003), 234-235.

[51] Jim Boyce, *Absolute Beginner's Guide to Microsoft Office ....*, 234.

The word appeared in red in the "Not in Dictionary list", and Ms. Word offers suggested changes in the "Suggestions list". The followings are buttons which can be used to accept or reject the change[52]:

1) Ignore Once: Leave this word alone (that is, reject the correction) and continue to the next misspelled word.
2) Ignore All: Don't change any occurrence of this word in the document.
3) Add To Dictionary: Add the unrecognized word to the dictionary so it won't be considered misspelled in the future.
4) Change: Change the spelling to the selected suggestion.
5) Change All: Change all occurrences of the word in the document to the selected suggestion.
6) AutoCorrect: Add the word and its correction to the AutoCorrect list. You should use this option if you frequently misspell or mistype the word in the same way.

(b) Checking Grammar

Ms.Word also checks your grammar. When it finds a word or phrase or sentence is grammatically incorrect, Ms.Word underlines the word or phrase or sentence with a green squiggly line. In checking grammar may identified phrase immediately click the word or phrase or sentence which are underlined with a green

---

[52] Jim Boyce, *Absolute Beginner's Guide to Microsoft Office ....*, 234.

squiggly line with the right button of the mouse, for example Ms.Word use to open the menus, shown in figure 2.5.[53]

| uses |
| Ignore Once |
| Grammar... |
| About This Sentence |
| Look Up... |
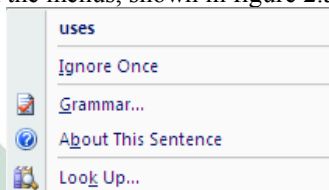
Figure 2.5 Menu
Grammar Checking

The followings are buttons which can be used to accept or reject the change shown in figure 2.5.

1) Uses: Showing the suggested word. The word "uses" appeared in top of the menu is the suggested changes offered by Ms.Word to fix the wrong grammar structure.

2) Ignore Once: Leave this word alone (that is, reject the correction) and continue to the next grammatically incorrect word.

3) Grammar: Showing the suggested word with the its sentence.

4) About This Sentence: Ms. Word will explain why this word is wrong. This can help you determine whether you want to change or ignore it.

5) Look Up: The detail explanation about the word by using the reference of Ms. Word's dictionary.

(2) Using the Dictionary

The use of the dictionary to check the spelling by looking the word up on it. It is

---

[53] Jim Boyce, *Absolute Beginner's Guide to Microsoft Office ....*, 238.

helpful to keep a list of words that have misspelled so these words can be learnt. Looking at words "A" dictionary not only tells how to spell a word, but also tells what part of speech the word is. Sometimes the word appears more than once, since it has different meanings and can be used as a different part of speech.[54]

3) Minimize reading time.

The stem usually contains those words or phrases which would otherwise have to be repeated in each option.[55] If test items are long and require extensive reading, administration time will be longer, because it takes time. Verbose items limit the number of items students can response (answer) in a set time, which has a negative effect on the reliability of test scores.

Paragraph will usually consist of the "standard" 100 – 200 words and the period is one of the writer's best friends. In most cases, when a sentence has 20 words or more, the writer starts using the way to end it. To save time, a sentence average is probably around 20 words. It means that every sentences will have the shortest sentences (as short as one word), medium-length sentences (around 12 to 18 words), and the longest sentences (around 18 to 25 words).[56]

To save reading time, the passage should be easy to read. How much easier is the graded of the passage by applying the formula to find that the graded passage has a "reading ease" score or not. If it has reading score around 60 which means that the graded of this passage was properly designed for the average high-school student.[57] The figure 2.6 which is written

---

[54] Marion Field, *Improve your Written English, Fifth-Edition* (Begbroke: How to Content, 2009), 63.

[55] J.B. Heaton, *Writing English Language Tests* ...., 30.

[56] Paula LaRocque, *The Book on Writing: The Ultimate Guide to Writing Well* (Portland: Marion Street Press, 2003), 14.

[57] Rudolf Flesch, *The Art of Readable Writing* (Evanston: Harper & Row Publisher, 1949), 148.

in the book *The Art of Readable Writing*, shows the estimates of these reading ease-score.[58]

Description of Style	Average Sentence Length	Average No. of Syll. per 100 Wds.	Reading Ease Score	Estimated Reading Grade
Very Easy	8 or less	123 or less	90 to 100	5th grade
Easy	11	131	80 to 90	6th grade
Fairly Easy	14	139	70 to 80	7th grade
Standard	17	147	60 to 70	8th and 9th grade
Fairly Difficult	21	155	50 to 60	10th to 12th grade (high school)
Difficult	25	167	30 to 50	13th to 16th grade (college)
Very Difficult	29 or more	192 or more	0 to 30	college graduate

Figure 2.6 Reading Ease-Score

It can be seen in figure 2.6 passage with a "reading ease" score around 60 is for 8th or 9th grade reading. The average high-school student will find it easy; they will be able to learn something from it without being bothered by reading difficulty which may takes their time. It can be seen in the figure 2.6. that a score of 45 means that the passage is college-level reading. A high-school students have to work more slowly (more takes their time) by reading around 40 words each sentence.[59]

Therefore, items writer should avoid excessive writing that uses too many words, because it can take a long time for the students to read.

b. Writing the Stem

1) Should test single language area (vocabulary and sentence structure should be measured in testing reading skill).

The stem should provide a complete idea of the knowledge to be indicated by choosing the right answer.[60] It may contain language areas: grammar or vocabulary (concerns on word meaning, word formation, and collocation).[61]

---

[58] Rudolf Flesch, *The Art of Readable Writing* ...., 148.
[59] Rudolf Flesch, *The Art of Readable Writing* ...., 149.
[60] Thomas. M. Haladyna, *Developing and Validating Multiple-Choice Test* ...., 68.
[61] J.B. Heaton, *Writing English Language Tests* ...., 9.

a) Vocabulary Test (concerned with word meaning, word formation and collocation).

A test of vocabulary measures students' knowledge of the meaning of certain words as well as the patterns and collocations in which they occured.[62] In figure 2.7 which is written in the book *Writing English Language Tests*, shows item which students are instructed to circle the letter at the side of the word which best completes the sentence.[63]

```
Did you . . . . . . . . . . that book from the school library?
A.  beg
B.  borrow
C.  hire
D.  lend
E.  ask
```

**Figure 2.7 Word Formation Type
to Test Vocabulary Testing Area**

In another common type of Vocabulary test students are given a passage to read and required to replace certain words listed at the end of the passage with their equivalent in the passage.[64]

Item used passage is required the students to choose the correct answer which is true according to the information stated in each sentence of passage. This particular knowledge is necessary for the understanding of the sentence.[65]

b) Grammar Test and Usage

These tests measure students' ability to recognise appropriate grammatical forms and to manipulate structures.[66] It is important to the students in mastering the grammatical area of the language they are learning. Thus, classroom tests

---

[62] J.B. Heaton, *Writing English Language Tests* ...., 9.
[63] J.B. Heaton, *Writing English Language Tests* ...., 9.
[64] J.B. Heaton, *Writing English Language Tests* ...., 9.
[65] J.B. Heaton, *Writing English Language Tests* ...., 35.
[66] J.B. Heaton, *Writing English Language Tests* ...., 9.

of grammar and usage can play a useful part in a language program.[67]

The type of multiple choice item preferred by many writers of Grammar test is the incomplete sentence type, with a choice of four or five options.[68] To test Grammar, use the options which contain the correct answer as part of the sentence which it can be compared on equal terms with the other options (distractors).[69]

A common grammar test is testing tenses. It is verbs changing form to indicate tense, or time that an action. English has six tenses: present, present perfect, past, past perfect, future, and future perfect. Each of the six tenses has a progressive form indicating a continuing action. Shifting in tense should occur when the time of the action changes.[70] Therefore, when testing tenses in multiple choice items, it should consistent with using time signals or it can be called as signal words. So, the students may know which form of verb change indicates the appropriate tenses which have those signal words.

Signal words help the students to put in the correct tense. The students can easily put in the correct form of the verb if the students know the signal word and which tense it demands. Table 2.1 shows some signal words in the following[71]:

---

[67] J.B. Heaton, *Writing English Language Tests* ...., 34.
[68] J.B. Heaton, *Writing English Language Tests* ...., 35.
[69] J.B. Heaton, *Writing English Language Tests* ...., 35.
[70] Mary K. McCaskill, *Grammar, Punctuation, and Capitalization: A Handbook for Technical Writers and Editors* (Hampton: Langley Research Center, 1998), 7.
[71] Mary K. McCaskill, *Grammar, Punctuation, and Capitalization* ....., 8.

Table 2.1 Signal Words for Tenses

Signal Words		Tenses
(1) every ...(day) (2)often	(3) seldom (4) sometimes (5) usually	Simple Present
(1) at the moment (2)Look!	(3) Listen! (4) Now	Present Progressive
(1) ... ago (2)last ...	(3) in 2010 (4) yesterday	Simple Past
(1) already (2)ever (3) so far	(4) up to now (5) yet	Present Perfect

Adopted by Mary K. McCaskill in *"Grammar, Punctuation, and Capitalization (A Handbook for Technical Writers and Editors)"*

In testing grammar may also require the students to choose the options which are true according to a particular syntax to test the understanding of the sentences.

2) Make directions as clear as possible.

The directions of stem clearly present the point to be addressed to tell the students what to do.[72] The primary purpose of the stem is to present the problem clearly and concisely. The students should be able to obtain the information what to do from the direction, read the stem and then answer required.[73] Therefore, the stem of multiple choice items should be written a way that the students know immediately what the focus of the items is.

3) Make the stem as brief as possible.

Test writer should make the stem briefly.[74] State stems as simply as possible.[75] Because the primary

---

[72] Albert Oosterhof, *Developing and Using Classroom Assessment* ...., 89.

[73] J.B. Heaton, *Writing English Language Tests* ...., 30.

[74] J.B. Heaton, *Writing English Language Tests* ...., 108.

[75] H. Douglas. Brown, *Language Assessment Principles and Classroom Practice* ...., 57.

purpose of the stem is to present the problem clearly and concisely.

A common mistake as making problem focused which is stated in the option and making a short and meaningless stem. This matter may make the students difficult to understand the item intended after reading the stem. We must include the problem or the central idea of the item in the stem rather than in the option to write focused item. When writing of the stem as brief as possible the students can get the point in the stem and let them choose among options.

4) Avoid irrelevant information (window dressing).

The stem should not contain extraneous information or irrelevant clues, it may confuse the problem being tested.[76] Because some items contain words, phrases, or entire sentences that have nothing to do with the problem stated in the stem, and vice versa, does not contain words, phrases, or sentences that make the students confused about what focus to do with the stem.[77]

Therefore, it is called window dressing if the sentence is used in the stem where the information is useless, without any purpose or value and vise versa. If the sentence is used in the stem without useful information needed, it does not have the focus about the problem and what to do to answer the items.

5) Avoid negative words in the stem.

Tamir cited research from the linguistic literature that negative words in the stem need about twice more working memory than positive words form of the same item. According to the source of evidence about negative items, it seems reasonable that negative wording should not be used in stems or in options.[78]

If a negative word is used, it should be stressed by placing it in bold type, capitalizing it, or

---

[76] J.B. Heaton, *Writing English Language Tests* ...., 30.
[77] J.B. Heaton, *Writing English Language Tests* ...., 108.
[78] Thomas. M. Haladyna, *Developing and Validating Multiple-Choice Test* ...., 111.

underlining it, or all of these. The reason is that the student often reads the negative word such as, "NOT/EXCEPT" and forgets to reverse the logic of the relation being tested. Although this is a popular format and it may perform sufficiently, this kind of item puts additional pressure on students in terms of working short-term memory. This is why the use of negative word is not recommended for item stems.[79]

c. Writing the Options

1) Use four options

Use As many options as possible, but four seems to be a natural limit. As well as the rule of writing multiple choice items in Indonesia, the number of answer options for Elementary and Junior High School are four optinos (A, B, D, C). The number of answer options for Senior High School and equivalent are five options (A, B, C, D, E).[80]

Some researches support the use of four options for conventional multiple choice items. According to Andres & del Castillo, Bruno & Dirkzwager, Haladyna & Downing, Landrum, Cashin, &Theis, Lord, Rodriguez, Rogers & Harley; Sax &Reiter, and Trevisan, Sax, & Michael summarize their research on the optimal number of options, evidence suggest that using more options per test item have a little advantage if each distractor is discriminating.[81] Thus, it should discriminate between the upper and lower group using the correct options.[82]

It means that every distractor must be able to trap the students, especially the students who did not master the items being tested, while clever students who are more knowledgeable about the material should not be trapped by distractors. Nevertheless,

---

[79] Thomas. M. Haladyna, *Developing and Validating Multiple-Choice Test* ...., 112.

[80] Pusat Penilaian Pendidikan, Kemendikbud, *Panduan Penulisan Soal Pilihan Ganda* (Jakarta: Kemendikbud, 2017), 14.

[81] Thomas. M. Haladyna, *Developing and Validating Multiple-Choice Test* ...., 112.

[82] M. David Miller, Robert L. Linn, *Measurement and Assessment in Teaching, Tenth-Edition* (New Jersey : Pearson Education, Inc.,Upper Saddle River, 2009), 357.

Haladyna and Downing found that many distractors do not discriminate.[83]

Therefore, the use of many options is possible, but four seems to be a natural limit. The optimum number of options for each multiple choice item is four in most public tests.[84]

2) Vary the location of the corret answer according to the number of options.

Assign the position of the right answer randomly.[85] The tendency of putting the correct answer in the same location of options makes the students to guess. Students may look for clues that will help them to guess. If the first option is usually the correct answer, the student will find this pattern and when in doubt choose A. Regarding this problem, often in placing the answer key, the teacher takes shortcuts by making certain patterns to make it easier in the process of correction. It will provoke students to follow the pattern that has been found to answer the items, with no belief in the truth of the answer.

To vary the location of the correct answer, the teacher should distribute the correct answer randomly in balanced proportions. In the process of determining answers key randomly and proportionally. There are two ways that may help in determining the proportion of the answers of multiple choice item, those are manual way and computerized way (using Ms.Excel). the manual way is proposed by Prof. Dr. H. Dja'far Shidiq, MA (Professor of IAIN North Sumatra) and the computerized way proposed by Dr. H. Halfian Lubis, SH., MA. (Head of PAI Development Curriculum section for Vocational High School).[86]

---

[83] Thomas. M. Haladyna, *Developing and Validating Multiple-Choice Test* ...., 112.
[84] J.B. Heaton, *Writing English Language Tests* ...., 28.
[85] Thomas. M. Haladyna, *Developing and Validating Multiple-Choice Test* ...., 113.
[86] Dja'far Shidiq - Halfian Lubis. "The Technique of Multiple Choice Answer Key Randomization" *Dunia adalah Ruang Belajar Sesungguhnya*, (https://rofiquez.wordpress.com, accessed on January 01, 2019)

The following are the description of manual and computerized way:

a) Manual Way

The following are the steps of this manual way. This manual way use the formula, the formula used is:

$$(\sum \text{Items} : \sum \text{option}) \pm 3$$

The example of this manual way described in the following:

The number of items made are 50 items with 4 options, those are A, B, C, D.
So it is known:
$\sum$ items: 50
$\sum$ options: 5

Then the proportion of the answer key is as follows:
(50 : 5) ± 3
10 ± 3

The calculation above means: one correct answer (key) cannot be more than 13 since (10+3) and not less than 7 since (10-3). You can determine the distribution of answer keys subjectively. The table 2.2 shows the example of answer key-distribution.

Table 2.2 Example of Answer Key-Distribution

Answer Key	Number of Items
A	10
B	11
C	9
D	13
E	7

Adopted by Dja'far Shidiq in "*The Technique of Multiple Choice Answer Key Randomization*"

The following are the informations of the table 2.2:
(1) Amounts other than maximum and minimum numbers (13 and 7) are odd. If even numbers occur it doesn't matter.
(2) The distribution of the number of keys on each option is not standard, so it depends on the items writer.
(3) The calculation of these proportions to avoid too much gap between the answer keys in each option.

By using This technique, the distribution of the number of answer keys on each option is quite equal with a number that is not too far.
b) Computerized Way (Using Ms.Excel)
The use of Ms.Excel is the second way to vary the location of the correct answer randomly in balanced proportions. The differences between the computerized way and the manual way are:
(1) In terms of key distribution in each item, the answer key distribution of the manual way is done manually, so the items writer can "set it up" subjectively. Whereas in the computerized way, the key distribution is automatically done randomly by the computer.
(2) In terms of the information about the number of answer keys in each option, in the manual

way this matter is known before the items writer spread the answer keys throughout the items number. In the computerized way, the number of answer keys in each option can be known after randomization is completed, by manually counting.

The following are the steps of this computerized way in using the Ms.Excel program to randomize the answer key:
(1) Open Ms.Excel
(2) Place the cursor in any column, for example, column "B3", then type this formula:

$$= RAND () * 10$$

Then press "enter"

(3) Place the cursor in the "B3" column, then press "F9". When press "F9" repeatedly will get different numbers.
(4) The number which appears, it shows the option. Only use number before comma. The following are the conversions of number to options are:

Number 1 = A,
Number 2 = B,
Number 3 = C,
Number 4 = D,
Number 5 = E,

The numbers used to determine which options (A, B, C, or D) are for each item number, are the numbers before comma, ignore the numbers that lie after the comma. When the number 0 (zero) or another number above number 5 appears, must press the "F9" key again to look for another number for the next item. The table 2.3 shows the example

of randomization number-result got from computerized way.

Table 2.3 Example of Randomization
Number-Result Gotten
from Computerized Way

Press of "F9"	Number which appears	Press again ?	Result	
			Items No.	Answer Key
	1,05707			
First Press	**3**,678701	No	1	C
Second Press	8,336431	Yes	2	B
Third Press	**2**,05616	No		
Fourth Press	6,75405	Yes	3	C
Fifth Press	**3**,079846	No		
Sixth Press	**5**,4711	No	4	E
Seventh Press	**2**,633902	No	5	B
Eighth Press	9,683978	Yes	6	B
Ninth Press	7,868893	Yes		
Tenth Press	**2**,48762	No		
....etc....				

Adopted by Halfian Lubis in "*The Technique of Multiple Choice Answer Key Randomization*"

The numbers used in the table 2.3 to determine the location of answer key in the options is a number that is "boldly written". The numbers that appear in different columns will be different even if the formula is the same.

The advantage of the computerized way (using Ms.Excel) is more accurate and scientific in randomization system. However when calculate the composition of the number of answer key, it might find a striking difference. It means that the difference of the number of answer key on each option is quite far. This can be dealt by using the manual way. Reduce the number of options which exceed 13 (10+3) then add the number of options which less than 7 (10-3).

It can be said that both ways of randomizing answer key can be used together. For randomization, use the computerized way (using Ms.Excel). After completing the randomization and to avoid the difference numbers of options that are quite far, control the number of keys on each option manually using manual way.

3) Place options in numerical order.

Answers should always be arranged in ascending or descending numerical order to unnecessarily increase the stress level for the test and we waste the student's time. The figure 2.8 which is written in the book *Developing and Validating Multiple-Choice Test Items (Third-Edition)*, shows two versions of the same item with the numerical order and non-numerical order.[87]

---

[87] Thomas. M. Haladyna, *Developing and Validating Multiple-Choice Test ....*, 113.

Wrong	Right
What is the cost of an item that normally sells for $9.99 that is discounted 25%?	What is the cost of an item that normally sells for $9.99 that is discounted 25%?
A. $5.00	A. $2.50
B. *$7.50	B. $5.00
C. $2.50	C. $6.66
D. $6.66	D. *$7.50

Figure 2.8 Numerical and Non-Numerical Order of Options

Therefore, options should always be arranged in numerical order (ascending or descending) to decrease the stress level and minimize taking time of the students.

4) Keep options independent; choices should not be overlapping.

If options are overlapping, these options are likely to give a clue to the students about the correct answer and the distractors. If a value contained in overlapping options is correct, the item may have two or more correct answers. The figure 2.9 which is written in the book *Developing and Validating Multiple-Choice Test Items (Third-Edition)*, illustrates this problem. Numerical problems that have ranges that are close make the item more difficult. More important in this example which is shown in the figure 2.9, Options A, B, C, and D overlap slightly. If the answer is age 25, one can argue that both C and D are correct though the writer of the item meant C. This careless error can be simply corrected by developing ranges that are distinctly different. The avoidance of overlapping options also will prevent embarrassing challenges to test items.[88]

---

[88] Thomas. M. Haladyna, *Developing and Validating Multiple-Choice Test ....*, 115.

> 22. What age range represents the physical "peak" of life?
>   A. 11 to 15 years of age
>   B. 13 to 19 years of age
>   C. 18 to 25 years of age
>   D. 24 to 32 years of age
>   E. over 32 years of age

Figure 2.9 Overlapping Options

Therefore, Make options independent, in the other word options should not be overlapping. If options are overlapping, these options are likely to give a clue to the students about the correct answer and the distractors. Also if a value contained in overlapping options is correct, the item may have two or more correct answers.

5) Keep the options homogeneous in content and grammatical structure.

The options should be parallel writing.[89] The use of options that are heterogeneous in content is often a clue to the student. Fuhrman suggested another way to view the issue of option homogeneity. If the correct answer is shorter or more specific or stated in other language, these tendencies might make the item easier.[90] Also keeping options homogeneous avoids the possibility of giving a clue to right answer. In the other word, only one feature at a time should be tested.[91] Grammatically error may clue the students to the right answer.[92]

If the options are heterogeneous in content, it may clue the students to choose the heterogeneous option.[93] Therefore, Keep the options homogeneous in content and grammatical structure. The use of options that are heterogeneous in content is often a cue to the student. A standard practice of keeping options

---

[89] Albert Oosterhof, *Developing and Using Classroom Assessment* ...., 90.
[90] Thomas. M. Haladyna, *Developing and Validating Multiple-Choice Test* ...., 116.
[91] J.B. Heaton, *Writing English Language Tests* ...., 29.
[92] Thomas. M. Haladyna, *Developing and Validating Multiple-Choice Test* ...., 119.
[93] Thomas. M. Haladyna, *Developing and Validating Multiple-Choice Test* ...., 119.

homogeneous avoids the possibility of giving information of a right answer.

6) Keep the length of options about the same.

One common fault happened in item writing is to make the sentence of correct answer longer than the distractors.[94] Since, it can give clues to the students by writing the sentence of correct answer longer than the distractors. So, it is necessary to make the correct option approximately the same length as the distractors.[95]

Therefore, keep the length of options about the same. The correct option should be written with the sentence thereabout same length as the distractors. Avoid one common fault in item writing that is to make the correct answer the longest.

7) "None of the above" should be used sparingly.

Using "none of the above" in some circumstances may force the student to solve the problem rather than choose the right answer. In these circumstances, the student may work backward, using the options to test a solution.[96]

Therefore, "none of the above" should be used sparingly. The equivalent options to 'none of the above' should be avoided.[97] In this case, the students may work backward, using the options to answer the item. It also may spend the time.

8) Avoid using "all of the above".

If student has partial information (knows that two of the three choices offered are correct), that information can clue the student into correctly choosing "all of the above". Since the purpose of a multiple choice test item is to test knowledge, using all of the above' seems to draw students into test-taking strategies more than directly testing for knowledge.[98]

[94] Thomas. M. Haladyna, *Developing and Validating Multiple-Choice Test ....*, 116.

[95] J.B. Heaton, *Writing English Language Tests ....*, 31.

[96] Thomas. M. Haladyna, *Developing and Validating Multiple-Choice Test ....*, 116.

[97] Albert Oosterhof, *Developing and Using Classroom Assessment ....*, 93.

[98] Thomas. M. Haladyna, *Developing and Validating Multiple-Choice Test ....*, 117.

Therefore, avoid using "all of the above". Because the purpose of a multiple choice item is to test knowledge, using "all of the above" may attract students into test-taking strategies more than directly testing for knowledge, in the other word it may give a clue into correctly choosing "all of the above". One alternative is simply avoid "all of the above" to ensure that there is one and only one right answer each multiple choice must only have one answer.[99] Then, make certain that the intended answer is clearly the only correct one.[100]

9) Avoid negative words such as "not" or "except".

The use of negatives such as 'not' and 'except' should be avoided in options as well as the stem. Sometimes, the use of these words in an item stem is unavoidable. In this situation, we should boldface, capitalize, italicize, or underline these words so that the students will not mistake the meaning of the item.[101]

Therefore, avoid negative words such as "not" or "except". In this case, if we use these kinds of words in options, we should bold face, capitalize, italicize, or underline these words.

10) Each option should be written in grammatically correct when placed in the stem.

Options should be written in grammatically correct when placed in the stem, except the items to test the specific grammar.[102] The figure 2.10 which is written in the book *Writing English Language Tests*, shows item 1 with stem ending with the determiner 'a', followed by options in the form of nouns phrases, when the answer (options) moved up to complete the stem, makes the sentence grammatically incorrect, then item 2 revised it.[103]

---

[99] J.B. Heaton, *Writing English Language Tests* ...., 28.
[100] H. Douglas. Brown, *Language Assessment Principles and Classroom Practice* ...., 58.
[101] Thomas. M. Haladyna, *Developing and Validating Multiple-Choice Test* ...., 117.
[102] J.B. Heaton, *Writing English Language Tests* ...., 29.
[103] J.B. Heaton, *Writing English Language Tests* ...., 29.

**Figure 2.10 Grammatically Incorrect Options
when Placed in the Stem and an Improved Version**

B. Review of Previous Study

Related to this research that focuses in analytical about testing, five studies are mentioned here as the previous studies:

The first study comes from Hana, under the title "*Analyzing the Multiple-Choice Items of Reading Test in Senior High School National Examination*."[104] This study aimed to discover whether the examination measured the objectives stated in School Based Curriculum (KTSP) including the items, the written texts, and the item types. The focus of this study is limited on the reading test of English subject. National Examination items are classified based on the cognitive process and the knowledge of the taxonomy theory which has been revised as proposed by Anderson. The study conducted by Hana has a similarity with the research conducted by the researcher, namely equally analyze multiple choice items. Meanwhile, the differences are this relevant study analyzes only reading content (reading test), and analyze paper based Final Examination. This study uses different theory and object of the study. Also this study did not analyze the construction of the test items.

The second study comes from Kim, under the title "*An Evaluation of Multiple Choice Test Questions Deliberately*

---

[104] Hana Dyah Merina, Undergraduate Thesis: "*Analyzing The Multiple-Choice Items of The Reading Test in Senior High School National Examination*" (Yogyakarta : Sanata Dharma University, 2009)

Designed to Include Multiple Correct Answers."[105] This study explored an alternative way of implementing multiple-choice questions with two or more correct answers by specifying in each question the number of answers examinees should select instead of using the traditional guideline. This study investigated the performance of three operational exams that use a standard multiple-choice format where the examinees are told how many answers they are to select. The result indicate that multiple choice items evaluated from these three exams performed at least as well as to the single-answer questions within the same exams. The study conducted by Kim has a similarity with the research conducted by the researcher, such as equally analyze multiple choice items. Meanwhile, the differences are this relevant study analyzes only multiple choice answer, analyze the effectiveness of two or more answers in multiple choice items experimentally, and analyze paper based test. This study uses different theory and object of the study. Also this study did not analyze the construction of the test items.

The third study comes from Poulomi, under the title *"Analysis of MCQs Item & Test Statistics from an assessment in a medical college of Kolkata, West Bengal".*[106] This study focus on how to asses item and test quality and to explore the relationship between difficulty index and discrimination index with the effectiveness of distractor. The study conducted by Poulomi Mukherjee has a similarity with the research conducted by the researcher, namely equally analyze multiple choice items. Meanwhile, the differences are this relevant study analyzes paper based test of medical college of Kolkata, West Bengal, and using numerical analysis. This study uses different theory and object of the study. Also this study did not analyze the construction of the test items.

---

[105] Kim Scott Thayn, Doctor of Philosophy Dissertation: "*An Evaluation of Multiple Choice Test Questions Deliberately Designed to Include Multiple Correct Answer*" (Provo : Brigham Young University, 2010), 62-63.

[106] Poulomi Mukherjee, "Analysis of MCQs Item & Test Statistics from an assessment in a medical college of Kolkata, West Bengal". *IOSR Journal of Dental and Medical Sciences (IOSR-JDMS).* Vol. 14, No.12, Dec. 2015, 47-52.

The fourth study comes from Surya, under the title "*Item analysis of multiple choice questions from an assessment of medical students in Bhubaneswar, India*".[107] The aim of this study is to evaluate the quality of MCQs, for creating a valid question bank for future use and to identify the low achievers, whose problems can be corrected by counselling or modifying learning methods. The study conducted by Surya has a similarity with the research conducted by the researcher, namely equally analyze multiple choice items. Meanwhile, the differences are this relevant study analyzes valid question bank, analyze paper based test of medical students in Bhubaneswar, India, and using numerical analysis. This study uses different theory and object of the study. Also this study did not analyze the construction of the test items.

The fifth study comes from Iffah Mursyidah Mayangsari, under the title "*An Analysis of UAS English Test of Second Semester 2008/2009 by Teacher-Made Test in SMA 2 Muhammadiyah Sidoarjo*".[108] This study aims to determine the quality of the final test items of English Test in 2008/2009 SMA 2 Muhammadiyah Sidoarjo. Seen in terms of content validity, reliability, difficulty index, and discrimination index. This study uses descriptive qualitative method. Then the result is checked with the criteria. The result is this test has high content validity and adequate reliability. But, theare are some items have unacceptable difficulty index. The study conducted by Iffah has a similarity with the research conducted by the researcher, namely using item analysis and analyzing english test. Meanwhile, the differences are this relevant research analyzes paper based final examination including validity, reliability, difficulty index, discrimination index qualitatively. This study uses different theory and object of the study. Also this study did not analyze the construction of the test items.

Seeing from the previous studies stated above which had been conducted before, the researcher concluded that all the previous studies have the simiarities and differences. Those previous studies

---

[107] Surya Kumar Namdeo, "Item analysis of multiple choice questions from an assessment of medical students in Bhubaneswar, India". *International Journal of Research in Medical Sciences.* Vol. 4, No. 5, May 2016, 1716-1719.

[108] Iffah Mursyidah Mayangsari, Undergraduate Thesis: "*An Analysis of UAS English Test of Second Semester 2008/2009 by Teacher-Made Test in SMA 2 Muhammadiyah Sidoarjo*" (Sidoarjo : UIN Sunan Ampel Surabaya, 2009).

could be the foundation of conducting this research. The previous studies mostly focusing on analyzing the options of multiple choice items, while this research focused on analyzing the construction of multiple choice items using the criteria of validated multiple choice items-construction.
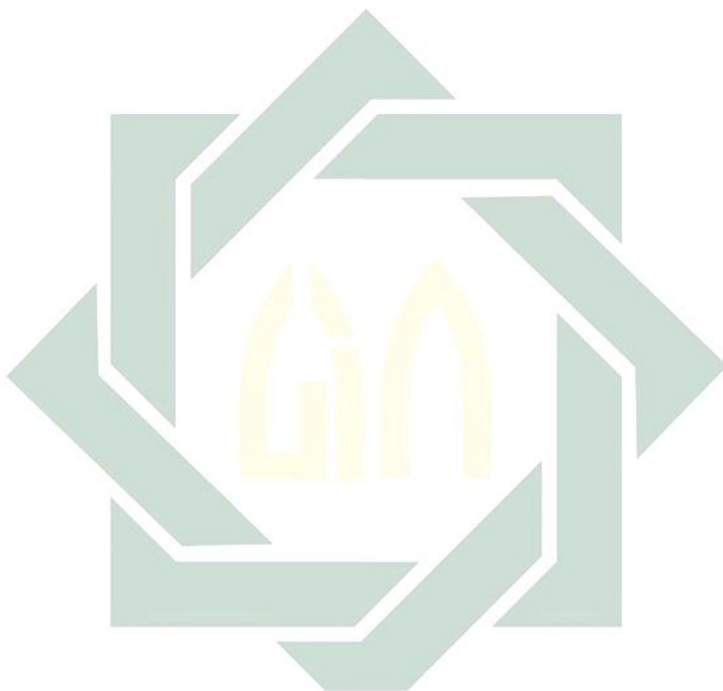
CHAPTER III
RESEARCH METHOD

This chapter describes the research method in analyzing the multiple choice items of Computer Based English Test used by SMP AlKhairiyah Surabaya, followed by research design, research location, data and source of data, research instruments, data collection technique, data analysis technique, checking validity, and research stages.

A. Research Design

Research design is the overall process as well as the strategy and tactics of science to guide the collection and analysis of data.[1] The method used in this research is qualitative analysis, particularly content analysis. The results of content analysis are described to obtain a comprehensive analysis in answering the research question.[2]

The researcher focuses on multiple choice items of Computer Based English Year-End Test conducted by SMP AlKhairiyah Surabaya at VIII grade in the 2017/2018 academic years. This Year-End Test (PAT) was conducted, because it has important role in measuring the achievement of students' learning outcomes at the end of second semester.

This research analyzed the construction of the item to determine its quality in term of whether the construction of multiple choice items taken from teacher made test fulfilled the criteria of validated multiple choice items-construction guidelines by Thomas M. Haladyna as the main theory or not, then revise the items.

B. Research Location

This research analyzed the multiple choice items of SMP AlKhairiyah Surabaya. It is a junior high school that is located on Jl. Sultan Iskandar Muda No.36, Ujung, Semampir, Kota SBY, East Java 60155. The researcher decided to choose this school as the object of this research based on the preliminary research found that SMP AlKhairiyah has some visions for their graduates, one of

---

[1] Paul S. Gray, *The Research Imagination: An Introduction to Qualitative and Quantitative Methods* (New York : Cambridge University, 2007), 34.

[2] John W. Creswell, *Research Design: Pendekatan Kualitatif, Kuantitatif, dan Mixed* (Yogyakarta: Pustaka Pelajar, 2012), 23.

45

which is mastering languages, one of languages is English.[3] It means that this school prioritizes their students in mastery English language. As one of effort to achieve this vision. This school makes the subject hour of English as the longest hour in learning and teaching process, also every morning the students gets vocabulary items from their English teacher.

In line with the prioritization of English subject in this school. The test made by the teacher to assess students can be prioritized too. In other word, the teacher is expected to make qualified items, but the English teacher at this school has never conducted item analysis to find out the quality of the test. The teacher only made multiple choice items for Year-End Test in accordance with all basic competence materials taught during the second semester.

The multiple choice items made by the teacher here are constructed without using the guidelines of how to construct multiple choice items.[4] It means that the quality of the multiple choice construction here have not been known whether it includes the multiple choice items that qualify as a good test construction or not. This situation can have an impact on misinterpretation of student achievement outcomes, because the error construction of the test items may make the students more difficult to response (answer) the item, not perform according to the expectations.[5] As it happened here, that the score of English Year-End Test (PAT) of some students of grade VIII is under the Minimum Exhaustiveness Criterion (KKM), that is 77.

Moreover, SMP AlKhairiyah Surabaya uses Computer Based Test (CBT) in various type of assessment. The most primary issues here is whether computer administration of a multiple choice item produces the same result as Paper Based Test (PBT) or not.[6]

Therefore, researcher tries to analyze the quality of multiple choice items of Computer Based English Year-End Test at grade VIII SMP AlKhairiyah Surabaya in the 2017/2018 academic years using validated multiple choice items-construction guidelines.

---

[3] *The Result of Interview with English Teacher of Grade IX - Mrs. Nyimas*, ...
[4] *The Result of Interview with English Teacher of Grade IX - Mrs. Nyimas*, ...
[5] J.B. Heaton, *Writing English Language Tests ....*, 37.
[6] Thomas. M. Haladyna, *Developing and Validating Multiple-Choice Test ....*, 93.

C. Data and Source of Data

1. Data

The researcher collected primary and secondary data to answer the research question. Those data will be described in the following:

a. Primary Data

The primary data used in this research were the data of Computer Based English Test (CBT for English). This primary data were obtained by collecting Computer Based English Test and analyzed it using criteria of validated multiple choice items-construction guidelines.

b. Secondary Data

The secondary data used in this research were some theories to support the data gained.

2. Source of Data

The sources of primary data were obtained from English teacher of grade VIII at SMP AlKhairiyah Surabaya. The primary data was obtained by analyzing the construction of multiple choice items of Computer Based English Test by using the criteria of validated multiple choice items-construction guidelines proposed by Thomas M. Haladyna (2004)[7] as the main theory to find how the construction of multiple choice items of Computer Based English Test is.

D. Research Instrument

This research instrument was used by the researcher to facilitate the researcher in collecting data, so that the analysis to undertake the findings of this research was easier to process. The researcher is the main instrument of this research, since this researcher used content analysis. As the main instrument, researcher was helped by the guidelines of the referencing systems based on the theory described in chapter 2. There are 20 criteria of the guidelines proposed by Thomas M. Haladyna (2004)[8] as the main theory. The 20 criteria including: (1) 5 criteria for multiple choice format-construction, (2) 4 criteria for multiple choice stem-construction, (3) 12 criteria for multiple choice options-construction. The guidelines were written in one table consist of

---

[7] Thomas. M. Haladyna, *Developing and Validating Multiple-Choice Test* ...., 98-121.
[8] Thomas. M. Haladyna, *Developing and Validating Multiple-Choice Test* ...., 98-121.

those criteria (3 categories: the format concerns, the writing of stem, the writing of options).

E. Data Collection Technique

Data collection in this research used documentation techniques, because the data were in the form of documents, including answer keys, and multiple choice items of Computer Based English Test. Data collection technique used by transferring the documents file taken from school party, the English teacher, on 23th November 2018, and treating the file as data to answer the research question. The researcher got 50 item numbers and answer keys.

F. Data Analysis Technique

In line with content analysis, the researcher analyzed the data descriptively. After getting the Computer Based English Test from the teacher, the researcher analyzed the multiple choice items-construction of Computer Based English Test using the criteria of validated multiple choice items-construction guidelines proposed by Thomas M. Haladyna (2004)[9]. The analysis result would show which criteria had been fulfilled and not by the multiple choice items-construction categories (format concerns, the writing of stem, and the writing of options) to gain the information about how the construction of multiple choice items of Computer Based English Test is how the construction of multiple choice items of Computer Based English Test is.

Finally, the researcher put the information from analysis result in findings of the research to be discussed. It would ensure whether the research question is answered or not.

G. Checking Validity

The researcher conducted the research to SMP AlKhairiyah Surabaya to get the multiple choice items of Computer Based English Test that has been designed by English teacher of grade VIII to meet the validity of the instrument. Then to find how the construction of multiple choice items of Computer Based English Test, whether they fulfilled the criteria or not, the multiple choice items-construction was analyzed based on guidelines that had been validated *(See appendix 1)*.The guidelines were validated by the expert of English language testing written on validated letter *(See*

---

[9] Thomas. M. Haladyna, *Developing and Validating Multiple-Choice Test ....*, 98-121.

appendix surat validasi). Then, to check the validity of the research findings, the researcher confirmed the findings with the theory used in this research.

H. Research Stages

1. Preliminary Research

Doing preliminary research is important to decide whether this research is can be done or not. The preliminary research had been done on 30th July 2018 doing interview with English teacher at SMP AlKhairiyah Surabaya. It gave great information related with the issue there that might be happened because of the error construction of multiple choice items on Computer Based English Test as described in the research background. Also preliminary research gave the information about for which school and for which class this research should be done as described in the research location. Besides, the researcher also did the library research to find the sources and which previous studies had similar areas with this research. Eventually, by doing this preliminary research, the researcher expected that this research should be done because of its importance.

2. Designing Investigation

In this stage, the researcher designed the investigation of how the construction of multiple choice items of Computer Based English Test is. To find whether each item fulfilled the validated multiple choice items-construction guidelines or not. These guidelines were written in the form of guidelines-table. The guidelines were proposed by Thomas M. Haladyna (2004)[10]. In term of getting the validity of this instrument, the researcher made validation sheet called validation letter *(See appendix "surat validasi")* which was validated by the expert of English language testing. These guidelines consist of 3 categories (multiple choice-format, multiple choice-stem, and multiple choice-options) which would be shown how the construction of multiple choice items of Computer Based English Test is based on which multiple choice items-construction categories fulfilled and not the criteria of guidelines *(See appendix 1)*.

---

[10] Thomas. M. Haladyna, *Developing and Validating Multiple-Choice Test ....*, 98-121.

3. Implementing Investigation

In this stage, the researcher began to collect the data from the English teacher at SMP AlKhairiyah Surabaya. The researcher asked the permission from the school party whether the researcher was allowed to do preliminary research and ask their Computer Based English Test or not. The researcher got the permission to do preliminary research with English teacher of grade VIII there. Then, the researher do some activities after doing preliminary research until finally get the Computer Based English Test file. The following are the description of these activities:

a. Firstly, the researcher got multiple choice items in the form of Ms.Excel table from the English teacher. Since the English teacher should make the items in the table before it is inputted to the CBT application. So, the teacher only have the items in the form of Ms.Excel table. The sample of the multiple choice items written in the form of Ms.Table which were researcher gotten, described in the following table:
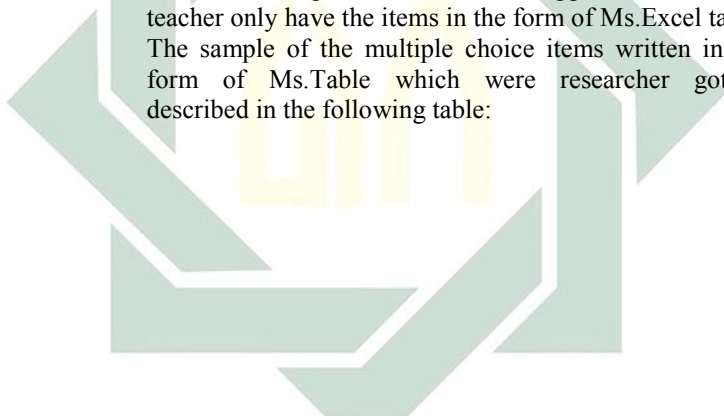
**Table 3.1 Multiple Choice Items of Computer Based
English Test in the Form of Ms.Excel Table**

N O	SOAL	JAWAB1	JAWAB2	JAWAB3	JAWAB4	JAWAB5	INI JAWAB	ACAK
1.	Complete these sentences using much or many number 1 to 5 Do you buy . . . food?	One	Much	Many	The		2	1
2.	we don't have . . . Bananas left	One	Much	Many	The		3	1
3.	.... (*See appendix 3*)	....	....	....	....	. . . .	....	...

As described in table 3.1, it can be seen that there are 9 columns in the top of this table. First column "NO" is for number of table, second column "SOAL" is for items (each item of Computer Based English Test is written here in sequence according to the number of the item), third column "JAWAB1" is for the first option of the item (if the options used in the multiple choice items are A, B, C, D, then the first option written in "JAWAB1" is option "A"), fourth column "JAWAB2" is for the second option which is option "B", fifth column "JAWAB3" is for the third option which is option "C",

sixth column "JAWAB4" is for the fourth option which is option "D", seventh column "JAWAB5" is for the fifth option (here does not need the fifth option, so this column is blank), eighth column "INI JAWAB" is the key answer of the item, ninth column "ACAK" is to randomize the key answer in each option.

b. Secondly, the multiple choice items which are in the form of Ms.Excel table should be inputted to CBT application. Since the items test which were tested to the students is in the form of CBT application, so it can not be printed as a paper sheet. So, the researcher had to take the screenshot of multiple choice items on Computer while the items was running on CBT application one by one item. The following are the display of multiple choice items of Computer Based English Test (*See appendix 4*).

c. Thirdly, after getting the Computer Based English Test, the researcher analyzed it using guidelines.

d. Finally, the analysis result would show which multiple choice items-construction categories fulfilled and not the criteria written in guidelines.

4. Analyzing data

In this stage, the researcher analyzed the data to get the answer of the research question as described in data analysis technique.

5. Concluding data

In this stage, to find research result, the researcher found the research final report by concluding the research finding.

This chapter presents the result of the research about the analysis of multiple choice items-construction on Computer Based English Test at SMP AlKhairiyah Surabaya.The researcher presents the results of the analysis in findings part and discuss the quality of multiple choice items-construction whether the items fulfill the criteria of validated multiple choice items-construction guidelines or not in terms of format, stem and options in discussion part.

A. Research Findings

The findings gotten in this research are by analyzing the construction of multiple choice items of Computer Based English Test consisted of 50 items using the criteria of validated multiple choice item-construction guidelines in terms of multiple choice format, stem and options.

The construction of Computer Based English Test at SMP AlKhairiyah Surabaya involves the activities in building the structure of multiple choice items in terms of designing format items, writing the stem and options. These activities were done by an English teacher.

In constructing the multiple choice items, must design the multiple choice format. Multiple choice format has parts: a stem and options (the correct choice, and several wrong answers, called distracters).[1] To construct the good multiple choice items, each of those components must fulfill the criteria of validated multiple choice items-construction guidelines.[2] The analysis result of the construction of multiple choice items which is categorized into 3 categories, those are the format concerns, the options-writing, and the stem-writing which were described in the following:

1. The Format Concerns of Multiple Choice Items-Construction

As described in chapter 2, a multiple choice-item is said to have good multiple choice format if the format have met 4 criteria of validated multiple choice items-format construction written in the guidelines for format concerns category. The

---

[1] Thomas. M. Haladyna, *Developing and Validating Multiple-Choice Test ....*, 68.
[2] Thomas. M. Haladyna, *Developing and Validating Multiple-Choice Test ....*, 97.

researcher used those guidelines to analyze whether each multiple choice-format of Computer Based English Test fulfills all of those criteria or not.

The analysis result shown, there are some items which the format fulfills each criteria and some do not fulfill the criteria. In short, the criteria of format concerns category covered by some items of Computer Based English Test are described in the following table:

Table 4.1 Criteria of Multiple Choice Format Covered by Computer Based English Test

No.	The Criteria of Format Concerns	
1.	Have stem and options	
2.	Items formatted vertically	
3.	Written with correct, punctuation, capitalization, and spelling.	
4.	Minimize Reading Time	

As described in table 4.1, there are 4 criteria covered by the format of Computer Based English Test. The following is an example of each item which fulfills and does not fulfill the format-criteria described in table 4.1:

e. Have stem and options

Based on the theories, a multiple choice-item is said to have good format-construction if the format have stem and options. The analysis result shown, all 50 items have stem and options. An example of these 50 items is described in the following:

**Figure 4.1 Example of Multiple Choice-Format
of Computer Based English Test
which has Stem and Options**

As displayed in figure 4.1 above, it can be seen that this item has a stem (as underlined by the red line) and options (as shown by the green right brace). It can be said that this item fulfills the criteria of validated multiple choice-format number 1 that is "have stem and options".

f.   Item formatted vertically

Based on the theories, a multiple choice-item is said to have good format-construction if the item formatted vertically. The analysis result shown, there are 42 items which have item formatted vertically. While there are 8 items which do not formatted vertically.

The researcher found that the writing of some items in this multiple choice items consist of a direction, stem, and options as well as these 42 items and they are formatted vertically between direction, stem, and options. As well as the items which do not have direction, they are structured by pilling up vertically. An example of these 42 items is described in the following:
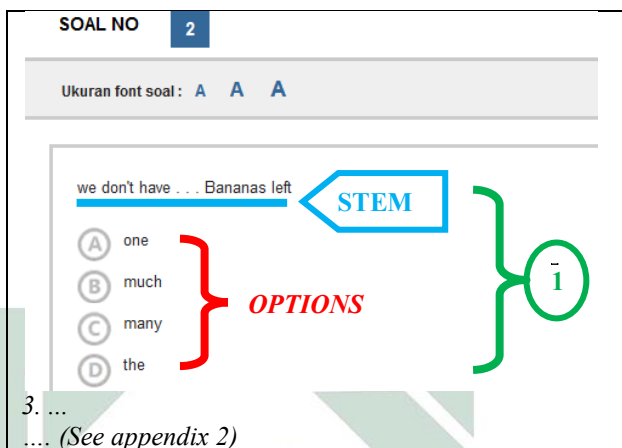
**Figure 4.2 Example of Multiple Choice Format
of Computer Based English Test Formatted Vertically**

As displayed in figure 4.2 above, it can be seen that this item has formatted vertically (as shown by the green right brace with number 1 beside it). The writing of the stem (as underlined by the blue line) and options (as shown by the red left brace) in this multiple choice-item are structured by pilling up vertically (not side by side). It can be said that this item fulfills the criteria of validated multiple choice-format number 2 that is "item formatted vertically".

There are 8 items are not formatted vertically. The researcher found that the writing of these 8 stems in this multiple choice items consist of direction, stem, and options, which between direction and stem are formatted horizontally, since they are structured side by side. An example of these 8 items is described in the following:

Figure 4.3 Example of Multiple Choice Format
of Computer Based English Test
Unformatted Vertically

As displayed in figure 4.3 above, it can be seen that the writing of this item consists of direction (as underlined by the blue line), stem (as underlined by the red line), and options (as shown by the green right brace), which between direction and stem are structured side by side (not formatted vertically). It can be said that this item does not fulfill the criteria of validated multiple choice-format number 2 that is "item formatted vertically".

g. Written with correct punctuation, capitalization, spelling & grammar

Based on the theories, a multiple choice-item is said to have good format-construction if the format written with correct punctuation, capitalization, and spelling. The analysis result shown, all 50 items are written with incorrect punctuation, capitalization, spelling and grammar. As the information, there are 43 items written with incorrect punctuation, 46 items written with incorrect capitalization, 6 items written with incorrect spelling, and 10 items written with incorrect grammar structure. In short, kinds of punctuation, capitalization, spelling, and grammar

which are covered by these 50 items are described in the following table:

Table 4.2 Kinds of Punctuation, Capitalization, Spelling, and Grammar
Covered by Computer Based English-Test

No.	Kinds of Punctuation, Capitalization, Spelling and Grammar
1.	Incorrect Punctuation
2.	Incorrect Capitalization
3.	Incorrect Spelling
4.	Incorrect Grammar

As described in table 4.2, there are four kinds of punctuation, capitalization, spelling and grammar covered in the items. The following is an example of each kind described in table 4.2:

1) Incorrect Punctuation

There are 43 items written with incorrect punctuation. An example of those 43 items is described in the following:

2. ....
.... (See appendix 2)

Figure 4.4 Example of Multiple Choice Format
of Computer Based English Test Written
with Incorrect Punctuation

As displayed in figure 4.4 above, it can be seen that this item has incorrect punctuation (as pointed by the green arrow). There was no period "." in the last word of the sentence "5". "Do" is the first word of the sentence in the stem but there is no punctuation to separate between them as the last word and the first. So, this matter connects the direction and the stem of the item. It can be said that this item does not fulfill the criteria of validated multiple choice-format number 3 that is "written with correct punctuation, capitalization, and spelling".

2) Incorrect Capitalization

There are 46 items written with incorrect capitalization. An example of those 46 items written with incorrect capitalization is described in the following:
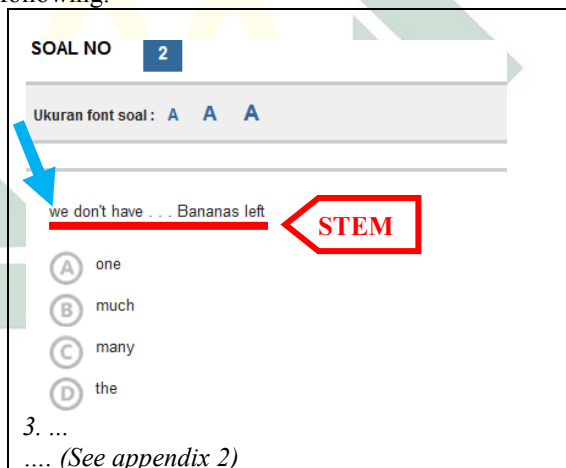
Figure 4.5 Example of Multiple Choice Format of Computer Based English Test Written with Incorrect Capitalization

As displayed in figure 4.5 above, it can be seen that this item has incorrect capitalization (as pointed by the blue arrow). It can be seen that the first letter of the word "we" is written in lowercase, whereas this

is the first word of the sentence, that is a word that must be written in uppercase. So, it can be said that this item does not fulfill the criteria of validated multiple choice-format number 3 that is "written with correct punctuation, capitalization, and spelling".

3) Incorrect Spelling

There are 6 items written with incorrect spelling. An example of those 6 items is described in the following:

SOAL NO **17**

Ukuran font soal : A **A** **A**

you/ listening/ to me

(A) you are listening to me

(B) you is listening to me

(C) you were listening to me

(D) you wasw listening to me

18. ...
22. ...
.... (See appendix 2)
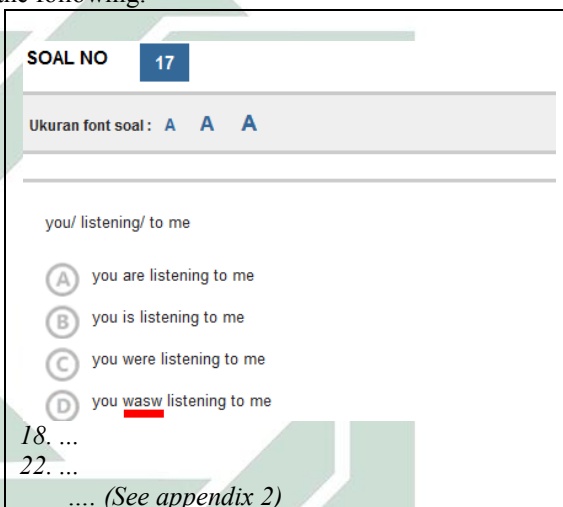
Figure 4.6 Example of Multiple Choice Format
of Computer Based English Test Written
with Incorrect Spelling

As displayed in figure 4.6 above, it can be seen that the item has incorrect spelling (as underlined by the red line). It shown that the word "wasw" is incorrect spelling. The word "wasw" supposed written to be "was". It can be said that this item does not fulfill the criteria of validated multiple choice-format number 3 that is "written with correct punctuation, capitalization, and spelling".

4) Incorrect Grammar

    There are 10 items written with incorrect grammar. An example of those 10 items is described in the following:

Figure 4.7 Example of Multiple Choice Format of Computer Based English Test Written with Incorrect Grammar

    As displayed in figure 4.7 above, it can be seen that the item has incorrect grammar (as shown by the red arrow). As we can see that the word "biggest" is incorrect grammar. The word "biggest" supposed written to be "the biggest". It can be said that this item does not fulfill the criteria of validated multiple choice-format number 3 that is "written with correct punctuation, capitalization, and spelling".

h. Minimize Reading Time

    Based on the theories, a multiple choice-item is said to have good format-construction if the item has been easy to read and it saves reading time (minimize reading time). To save reading time, a sentence average is probably around 20 words. The shortest sentence has one word, the medium-length sentence has around 12 to 18 words, and the longest sentence has around 18 to 25 words. The analysis result shown, all 50 items have been easy to read.

The researcher found that the writing of these 50 multiple choice items only consist of simple sentence. While the items written with the passage consist of a sentence average which is probably around 20 words. An example of these 50 items had brief stem with the longest sentence is described in the following:

> *36. Read the following text to answer question 36 to 40*
> *It was Saturday, august 23rd the weather was good. It wasn't rainy. The sun was shining our school was holiday a special event. It was ''performance day''. The event started at 9 a.m. every student was taking part in the event. Some students were playing chess and some others were taking part in face painting. I was working with my team. We were doing an experiment. The ''performance day'' finished at 2 p.m. we were all happy. We had good time. What is the text about?*
> A. *Descriptive Text*
> B. *Recount text*
> C. *Long text*
> D. *Story*
> *41. ....*
> *.... (See appendix 2)*

STEM has passage

STEM in the form of quesion

Figure 4.8 Example of Multiple Choice-Stem of Computer Based English Test Minimized Reading Time

As displayed in figure 4.8 above, it can be seen that the writing of this item consists of a stem which is written in the form of question (as underlined by the light blue line), and the stem has passage (as underlined by colorful line) which has a sentence average around 20 words. The first sentence has 5 words (as underlined by the red line), the second sentence has 4 words (as underlined by the dark blue line), the third sentence has 3 words (as underlined by the brown line), the fourth sentence has 11 words (as underlined by the purple line), the fifth sentence has 4 words (as underlined the yellow line), the sixth sentence has 6 words (as underlined by the orange line), the seventh

sentence has 8 words (as underlined by the aqua line) the eighth sentence has 14 words (as underlined by the green line), the ninth sentence has 6 words (as underlined by the gray line), the tenth sentence has 5 words (as underlined by the black line), the eleventh sentence has 7 words (as underlined by the pink line), the twelfth sentence has 4 words (as underlined by the light green line), and the thirteenth sentence has 4 words (as underlined by the dark red line). In the other word, this stem has passage which is written with under 25 words (the longest standard sentences) per-sentence. It can be said that this item fulfills the criteria of validated multiple choice-format number 4 that is "minimize reading time".

2. **The Stem-Writing of Multiple Choice Items-Construction**

As described in chapter 2, a multiple choice-item is said to have good multiple choice stem if the stem have met 5 criteria of validated multiple choice stem-construction written in the guidelines for the writing of stem category. The researcher used those guidelines to analyze whether each multiple choice-stem of Computer Based English Test fulfills all of those criteria or not.

The analysis result shown, there are some items which the stem-construction fulfill each criteria and some do not fulfill the criteria. In short, the criteria of stem-writing category covered by some items of Computer Based English Test are described in the following table:

Table 4.3 Criteria of Multiple Choice-Stem
Covered by Computer Based English Test

No.	The Criteria of Writing of Stem
1.	Should test single language area (vocabulary and sentence structure should be measured in testing reading skill)
2.	Have clear direction.
3.	Have brief stem.
4.	Avoid irrelevant information (window dressing).
5.	Written without negative word in the stem "NOT or EXCEPT" (If a negative word is used, it should be stressed by placing it in bold type, capitalizing it, or underlining it, or all of these).

As described in table 4.3 above, there are five criteria covered by the stem-writing of Computer Based English Test. The followings are some examples of items which fulfill and do not fulfill the stem-writing criteria described in table 4.3:

a. Should test single language area (Vocabulary and sentence structure (Grammar) should be measured in testing reading skill).

Based on the theories, a multiple choice-item is said to have good stem-construction if the stem tests single language area (Vocabulary area or Grammar area). The researcher found that the writing of stems in this multiple choice items of Computer Based English Test focused on reading test, which from 50 items, 45 items test single language area while 5 items do not test single language area.

The researcher found that from these 45 items, 10 items are written to test vocabulary area, 35 items are written to test grammar area. In short, the single language area testing covered by these 45 items are described in the following table:

Table 4.4 Single Language Area Testing
Covered by Computer Based English-Test

No.	The Single Language Area	
1.	Test vocabulary	
2.	Test grammar	

As described in table 4.4, there are two kinds of stem which test single language area. The following is an example of each kind of stem described in table 4.4:

1) Test vocabulary

There are 10 items written to test vocabulary area. An example of these 10 items is described in the following:

36. Read the following text to answer question 36 to 40 *It was Saturday, august 23rd the weather was good. It wasn't rainy. The sun was shining our school was holiday a special event. It was ''performance day''. The event started at 9 a.m. every student was taking part in the event. Some students were playing chess and some others were taking part in face painting. I was working with my team. We were doing an experiment. The ''performance day'' finished at 2 p.m. we were all happy. We had good time. What is the text about?*

A. *descriptive test*
B. *recount text*
C. *long text*
D. *story*

37. ...

.... (See appendix 2)

STEM has passage

STEM in the form of question

OPTIONS

Figure 4.9 Example of Multiple Choice Stem of Computer Based English Test Tested Vocabulary Area

As displayed in figure 4.9 above, it can be seen that the item consists of stem which has passage (as underlined by the blue line) and written in the form of question. The item also has the options (as shown by the red right brace) which consists of sentence with different vocabulary items. As described in chapter 2, said the characteristics of Vocabulary test are written in passage and used different vocabulary items. This item is written only to measure the depth of *vocabulary* knowledge of the students. It shown that this item tests single language area, that is Vocabulary test. It can be said that this item fulfills the criteria of validated multiple choice-stem number 1 that is "should test single language area".

2) Test grammar

There are 35 items written to test grammar area. As the information, these 35 multiple choice items test 3 different Grammar elements, those are 10 items test

Quantifier, 20 items test Verb Tenses, and 5 items test Degree of Comparison. In short, grammar area testing covered by these 35 items described in the following table:

Table 4.5 Grammar Elements
Covered by Computer Based English-Test

No.	Elements of Grammar
1.	Quantifier
2.	Verb Tenses
3.	Degree of Comparison

As described in table 4.5, there are 3 elements of grammar covered by grammar area testing. The following is an example of each element described in table 4.5:

a) Quantifier

There are 10 items which test Quantifier. An example of these 10 items is described in the following:

Figure 4.10 Example of Multiple Choice-Stem of Computer Based English Test Tested Quantifier in Grammar Area

As displayed in figure 4.10 above, it can be seen that the stem written with incomplete sentence (as underlined by the blue line). The options (as shown by the red right brace) contain of sentence with different grammar structures which include Quantifier, since the options use quantifier words, those are "there are", "there were", "there is", and "there was" Here, the students were asked to choose the option which Quantifier "there..." is proper to complete the sentence in the stem then put their choice in the gap. It means that this item exists to test the Quantifier structure knowledge of the students. This item tests single language area that is Grammar testing. It can be said that this item fulfills the criteria of validated multiple choice-stem number 1 that is "should test single language area".

b) Verb Tenses

There are 20 items which test Verb Tenses. An example of these 20 items is described in the following:

Figure 4.11 Example of Multiple Choice-Stem of Computer Based English Test Tested Verb Tenses in Grammar Area

As displayed in figure 4.11 above, it can be seen that the stem written with incomplete sentence (as underlined by the blue line). The options (as shown by the red right brace) contain of sentence with different grammar structures which include Verb Tenses, since the options use different changes of verbs form. Here, the students were asked to choose the option which verb tense is proper to complete the sentence in the stem then put their choice in the gap. It means that this item exists to test the verb tenses structure knowledge of the students. So, this item tests single language area that is Grammar testing. It can be said that this item fulfills the criteria of validated multiple choice-stem number 1 that is "should test single language area".

c) Degree of Comparison

There are 5 items test Degree of Comparison. An example of these 5 items is described in the following:

21. ....
.... (See appendix 2)

Figure 4.12 Example of Multiple Choice-Stem of Computer Based English Test Tested Degree of Comparison in Grammar Area

As displayed in figure 4.12 above, it can be seen that the stem written with incomplete sentence (as underlined by the blue line). The options (as shown by the red right brace) contain of sentence with different grammar structures which include Degree of Comparison, since the options use Degree of Comparison words, one of which is "more..." Here, the students were asked to choose the option which Degree of Comparison word is proper to complete the sentence in the stem then put their choice in the gap. It means that this item exists to test the Degree of Comparison structure knowledge of the students. This item tests single language area that is Grammar testing. It can be said that this item fulfills the criteria of validated multiple choice-stem number 1 that is "should test single language area".

There are 5 items which do not test single language area. These 5 stems test vocabulary and Grammar testing area. The following is an example of these 5 items:
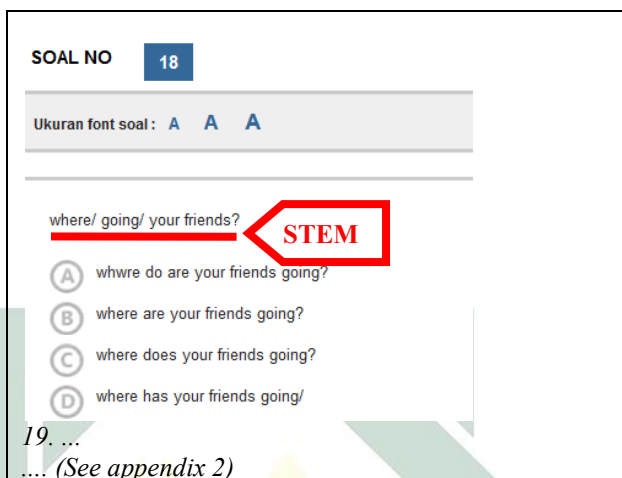
**Figure. 4.13 Example of Multiple Choice-Stem
of Computer Based English Test Tested
Two Language Areas**

As displayed in figure 4.13 above, it can be seen that the stem (underlined by the red line) written with word formation. Here, the students were asked to arrange the stem-sentence in correct order. As described in chapter 2, word formation is one of the types of item which tests Vocabulary area. Also it can be seen that the stem written with incomplete sentence. So, the students were asked to look for the missing word from the stem grammarly. It means that this item tests both Vocabulary area and Grammar area. It can be said that this item does not fulfill the criteria of validated multiple choice-stem number 1 that is "should test single language area".

b. Have clear direction

Based on the theories, a multiple choice-item is said to have good stem-construction if the stem has clear direction. The analysis result shown, from 50 items, only 1 item has clear direction while 49 items do not have clear direction.

The writing of stem in this 1 multiple choice item consists of a direction and stem sentence which are

formatted vertically. In the other word, this 1 item is structured by pilling up vertically. This item is described in the following:

> 42. *Read the following text to answer 41 to 45*
> *I want to tell you about my favourite uncle. His name is Imam. He is journalist at a tv station. He reports about the low and crimes, he works hard. In his spare time, he likes reading. His favourite book is about crimes. We often have time to discuss many things. He has good knowledge because he likes reading a lot.*
> *what is the text about?*
> A.   *recount text*
> B.   *descriptive text*
> C.   *long text*
> D.   *functional text*

STEM in the form of question

STEM has passage

**Figure 4.14 Example of Multiple Choice-Stem
of Computer Based English Test Written
with Clear Direction**

As displayed in figure 4.14 above, it can be seen that the item has clear direction. The stem-writing of this item consists of direction (as shown with the red line), stem (has passage (as underlined by the blue line)) and it is written in the form of question (as underlined by the orange line). Between direction and stem are structured by pilling up vertically. In the other word, this item has formatted vertically. It may make the direction as clear as possible. It can be said that this item fulfills the criteria of validated multiple choice-stem number 2 that is "have clear direction".

There are 49 items which do not have clear direction. As the information from these 49 items, 8 items are written with stem and direction connect each other (formatted horizontally) and 41 items do not have direction. In short, kinds of direction covered by these 49 items are described in the following table:

Table 4.6 Kinds of Unclear Direction
Covered by Computer Based English-Test

No.	Kinds of Unclear Direction
1.	Direction and stem connect each other (formatted horizontally
2.	Do not have direction

As described in table 4.6, there are 2 kinds of unclear direction covered in these multiple choice items. The following an example of each kind of unclear direction described in the table 4.6:

1) Direction and stem connect each other (formatted horizontally)

There are 8 items have direction, stem, and options, which between direction and stem are written with connect each other (formatted horizontally). The direction and stem of these 8 items are structured side by side. The following is an example of these 8 items:

Figure 4.15 Example of Multiple Choice-Stem
of Computer Based English Test
which does not have Clear Direction
(Direction and Stem Connect each Other)

As displayed in figure 4.15 above, the item does not have clear direction. The writing of this stem consists of direction (as shown with the blue line), stem (as shown with the red line) which between direction and stem sentence are structured side by side. It makes there are no differences between direction and stem. They look like as one sentence even though these are two sentences. It makes the direction unclear. It can be said that this item does not fulfill the criteria of validated multiple choice-stem number 2 that is "have clear direction".

2) Do not have direction

There are 41 items do not have direction. Without direction may make the students confused about what to do to answer the items. An example of these 41 items is described in the following:

Figure 4.16 Example of Multiple Choice-Stem of Computer Based English Test which does not have Clear Direction (without Direction)

As displayed in figure 4.16 above, it can be seen that the item does not have clear direction. The writing of this item does not have a direction. The item only

has stem (as underlined by the red line) and options (as shown by the blue right brace). Without direction may make the students confused about what to do to answer this item. It may make the direction unclear. It can be said that this item does not fulfill the criteria of validated multiple choice-stem number 2 that is "have clear direction".

c. Have brief stem

Based on the theories, a multiple choice-item is said to have good stem-construction if the stem written with brief stem. A stem is said to be brief stem if it has sentences or paragraph not longer than 100 words. The analysis result shown, all 50 items have brief stem. As the information, 47 items have short sentences, and 3 items have the longest sentences (passage).

There are 47 items which have short sentences in the stem with the length of the sentence are not longer than 10 words. The short sentences only have around 3-10 words. An example of these 47 items is described in the following:

Figure 4.17 Example of Multiple Choice-Stem of Computer Based English Test which has Brief Stem with the Short Sentence

As displayed in figure 4.17 above, it can be seen that the item has brief stem. The researcher found that the writing of this item consists of a stem which only has 4 words. In the other word, this item has stem with total words under 100. It can be said that this item fulfills the criteria of validated multiple choice-stem number 3 that is "have brief stem".

There are 3 items which have the longest sentences in the stem. The writing of these 3 items have the length of the longest sentence or paragraph not longer than 100 words. The longest paragraph only have 98 words, which is the stem written in the form of passage. The following is an example of these 3 items:

36. Read the following text to answer question 36 to 40
It was Saturday, august 23rd the weather was good. It wasn't rainy. The sun was shining our school was holiday a special event. It was '' performance day''. The event started at 9 a.m. every student was taking part in the event. Some students were playing chess and some others were taking part in face painting. I was working with my team. We were doing an experiment. The '' performance day'' finished at 2 p.m. we were all happy. We had good time. What is the text about?

 A. *Descriptive Text*
 B. *Recount text*
 C. *Long text*
 D. *Story*
41. ...
.... (See appendix 2)

STEM has passage

STEM in the form of question

Figure 4.18 Example of Multiple Choice-Stem of Computer Based English Test which has Brief Stem with the Longest Text

As displayed in figure 4.18 above, it can be seen that the item has brief stem. The writing of this item consists of a stem and it has a passage (as underlined by red line) which consists of 83 words, the total number under 100. It

can be said that this item fulfills the criteria of validated multiple choice-stem number 3 that is "have brief stem".

d. Avoid irrelevant information (window dressing)

Based on the theories, a multiple choice-item is said to have good stem-construction if the stem does not have irrelevant information or it can be called "window dressing". Stem is said to have window dressing if the item contains words or sentences that have nothing to do with the problem stated in the stem, and vice versa, does not contain words or sentences that may make the students confused about what focus to do to answer the stem. Related to this window dressing, some stems have direction, time signal for verb tenses, and text in passage-stem. In terms of those three, there are 7 stems are free from window dressing and there are 43 stems have window dressing. The stems in this 7 multiple choice items have direction and have text (passage) to test vocabulary. In short, kinds of these 7 stems which are free from window dressing covered by Computer Based English Test described in the following table:

Table 4.7 Kinds of Clear Direction
Covered by Computer Based English-Test

No.	Kinds of Stem which are Free from Window Dressing
1.	Have direction
2.	Have text (passage) to test vocabulary

As described in table 4.7, there are 2 kinds of stem which are free from window dressing. The following is an example of each kinds of stem described in the table 4.7:

1) Have direction

There are 7 items are free from window dressing and these 7 items have direction. Using direction may make the students focus about what to do to answer the items. The following is an example of these 7 items:

> 42. _Read the following text to answer 41 to 45_
>
> _I want to tell you about my favourite uncle. His name is Imam. He is journalist at a tv station. He reports about the low and crimes, he works hard. In his spare time, he likes reading. His favourite book is about crimes. We often have time to discuss many things. He has good knowledge because he likes reading a lot._
>
> _what is the text about?_
> A. _recount text_
> B. _descriptive text_
> C. _long text_
> D. _functional text_
>
> 36. ...
> .... _(See appendix 2)_

Figure 4.19 Example of Multiple Choice-Stem of Computer Based English Test which has Direction

As displayed in figure 4.19 above, it can be seen that the item has direction (as underlined by red line). By using the direction students may focus about what to do to answer the items. The direction written with the sentence "_Read the following text to answer 41 to 45_". So, the students may know that they have to read the text to answer the items number 41 to 45. It can be said that this item fulfills the criteria of validated multiple choice-stem number 4 that is "avoid window dressing".

2) Have text (passage) to test vocabulary

There are 7 items are free from window dressing. From these 7 items, 3 items have text (passage) to test vocabulary and 4 items do not have the text (passage), since they do not need the text (passage) to answer the items. in these 3 items, the options are given in certain words listed to choose the

correct answer which is true according to the information stated in each sentence of passage. The following is an example of these 3 items:

42. *Read the following text to answer 41 to 45*
I want to tell you about my favourite uncle. His name is Imam. He is journalist at a tv station. He reports about the low and crimes, he works hard. In his spare time, he likes reading. His favourite book is about crimes. We often have time to discuss many things. He has good knowledge because he likes reading a lot.

what is the text about?
    A. *recount text*
    B. *descriptive text*
    C. *long text*
    D. *functional text*

36. ...
41. ...
.... (See appendix 2)

STEM has passage

OPTIONS

Figure 4.20 Example of Multiple Choice-Stem of Computer Based English Test which has Text (Passage)

As displayed in figure 4.20 above, it can be seen that the stem of the item has passage. Here, the options (as shown by the red right brade) are given in the cerain words listed to choose the correct answer which is true according to the information stated in the passage as asked in the stem (as underlined by the green line). So, the studens can read the text (passage) written above the options to gain the information stated there to answer the items. It can be said that this item fulfills the criteria of validated multiple choice-stem number 4 that is "avoid irrelevant information (window dressing)".

There are 43 items have window dressing. From these 43 items, there are some items have window dressing because of the use of time signal to test Verb Tenses, the

use of direction, and the use of passage to test Vocabulary. In short, the window dressing covered by these 43 items described in the following table:

Table 4.8 Window Dressing
Covered by Computer Based English-Test

No.	Window Dressing
1.	The use of time signal for Grammar test
2.	The use of direction
3.	The use of text (passage)

As described in table 4.8, there are three kinds of window dressings. There some items have only one kind of window dressing, others have two window dressings, and the other have all three window dressing listed in the table 4.8. The following is an example of each item which has window dressings described in table 4.8:

1) The use of time signal for Grammar test

As described in chapter 2, when testing verb tenses in multiple choice items, it should consistent with using time signals or it can be called as signal word. So, the students may know which form of verb change indicates the appropriate verb tenses which have those signal words.

There are 43 items have window dressing. From these 43 items, 20 items test Verb Tenses. From these 20 items, 12 items do not have time signals in their stems and 8 items have time signals in their stems to test Verb Tenses.

There are 12 items do not have time signal, even they test Verb Tenses. The following is an example of these 12 items:
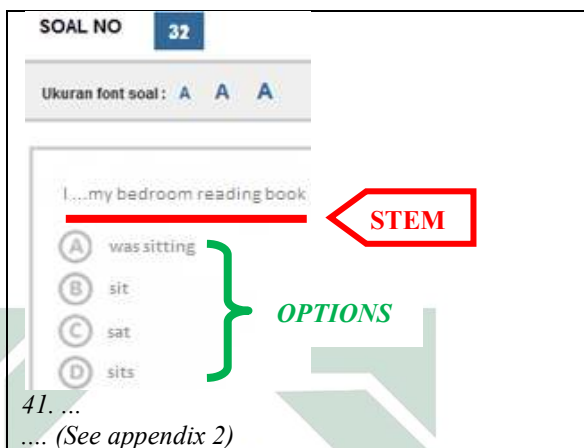
Figure 4.21 Example of Multiple Choice-Stem of Computer Based English Test which does not have Time Signal to test Verb Tenses in Grammar Area

As displayed in figure 4.21 above, it can be seen that the stem does not have time signal to test Verb Tenses. Even though, the options (as shown by the green right brace) contain of sentence with different grammar structures which include verb tenses, since the options use different changes of verbs form. Here, the students may be difficult to choose the option which verb changing form is proper with time in action to complete the sentence in the stem, since the stem does not have time signal. This may make the students not focus about what to do with the stem to answer the items, because of the lack of information stated in the stem. So, this item has window dressing. It can be said that this item does not fulfill the criteria of validated multiple choice-stem number 4 that is "avoid irrelevant information (window dressing)".

There are 8 items have time signals in their stems to test Verb Tenses. The following is an example of these 8 items:

**Figure 4.22 Example of Multiple Choice-Stem
of Computer Based English Test which has Time
Signal to test Verb Tenses in Grammar Area**

As displayed in figure 4.22 above, it can be seen that the stem has time signal to test Verb Tenses. Since, the stem has time signal "yesterday" (as underlined by the blue line) and the options (as shown by the green right brace) contain of sentence with different grammar structures which include verb tenses, since the options use different changes of verbs form. Unfortunately, this stem do not have direction. That may make the students may not focus about what to do with the stem to answer the item. Although this item has time signal, but this item does not have direction. So, this item has window dressing. It can be said that this item does not fulfill the criteria of validated multiple choice-stem number 4 that is "avoid irrelevant information (window dressing)".

2) The use of direction

There are 43 items have window dressing in their stems. From these 43 items, 41 items do not have direction and 2 items have direction in their stems.

There are 41 items do not have direction. The following is an example of these 41 items:

Figure 4.23 Example of Multiple Choice-Stem of Computer Based English Test which does not have Direction

As displayed in figure 4.23 above, it can be seen that the item does not have direction. The item only has a stem (as underlined by the red line) and options (as shown by the blue right brace). Without direction may make the students confused about what to do with the stem to answer this item. It can be said that this item does not fulfill the criteria of validated multiple choice-stem number 4 that is "avoid irrelevant information (window dressing)".

There are 2 items have direction in their stems. The following is an example of these 2 items:

**Figure 4.24 Example of Multiple Choice-Stem
of Computer Based English Test
which has Direction**

As displayed in figure 4.24 above, it can be seen that the item has direction (as underlined by red line). By using the direction written with the sentence "*Complete these sentences using present tense number 11 to 15"*, the students may know that they have to complete the text with using present tense to answer the items number 11 to 15. Even though, the options (as shown by the green right brace) contain of sentence with different grammar structures which include verb tenses, this stem does not have time signal. It may make difficult to choose the option which verb changing form is proper with time in action to complete the stem. Although this item has direction, but does not have time signal. So, this item has window dressing with the lacking information. It can be said that this item does not fulfill the criteria of validated multiple choice-stem number 4 that is "avoid window dressing".

3) The use of text (passage)

There are 43 items have window dressing. From these 43 items, 7 items test Vocabulary and these 7 items need text (passage) to answer the items, but these 7 items do not have passage.

The options of these 7 items are in certain words listed to choose the correct answer according to the passage, but there is no passage in the stem. The following is an example of these 7 items:
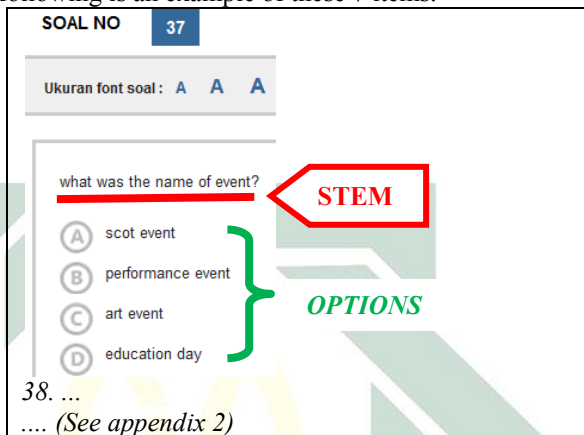
Figure 4.25 Example of Multiple Choice-Stem of Computer Based English Test which does not have Passage to Test Vocabulary Area

As displayed in figure 4.25 above, it can be seen that the stem does not have passage. Here, the options (as shown by the green right brace) are given in cerain words listed to choose the correct answer which is true according to the information stated in the passage as asked in the stem (as underlined by the red line). But the passage is not written above the options to ease the students in answering the items. This may make the studens difficult to read the passage in gaining the information stated there to answer the items. So, this item has window dressing. It can be said that this item does not fulfill the criteria of validated multiple choice-stem number 4 that is "avoid irrelevant information (window dressing)".

e. Written without using negative words in the stem (If a negative word is used, it should be stressed by placing it in bold type, capitalizing it, or underlining it, or all of these).

Based on the theories, a multiple choice-item is said to have good stem-construction if the stem does not use negative word or use by placing it in bold type, capitalizing it, or underlining it. The analysis result shown, all 50 items do not use negative word. It can be said that all of 50 items fulfills the criteria of validated multiple choice-stem number 5 that is "written without using negative words in the stem".

3. The Options-Writing of Multiple Choice Items-Construction

As described in chapter 2, a multiple choice-item is said to have good option-construction if the items have met 10 criteria of validated multiple choice items-options written in the guidelines. The researcher used those guidelines to analyze whether each multiple choice-option in Computer Based English Test fulfills all those criteria or not.

The analysis result shown, there are some items which the options fulfills each criteria and some do not fulfill the criteria. In short, the criteria of options-writing category covered by some items of Computer Based English Test are described in the following table:

Table 4.9 Criteria of Multiple Choice-Options
Covered by Computer Based English Test

No.	Criteria of writing the stem
1.	Use four options
2.	Have varied location of the right answer according to the number of options.
3.	Options written in numerical order (sequenced).
4.	Options written independently; choices should not be overlapping.
5.	Options homogeneously in content and grammatical structure.
6.	Have options with the same length.
7.	Written without using "None of the above".
8.	Written without using "all of the above" (one right answer).
9.	Written without using negative words in the options. (If a negative word is used, it should be stressed by placing it in bold type, capitalizing it, or underlining it, or all of these).
10.	Written each option in grammatically correct when placed in the stem (except the items to test the specific grammar).

As described in table 4.9, there are 12 criteria covered by the writing of options in Computer Based English Test. The following is an example of each criteria described in table 4.9 which is fulfilled and not by multiple choice items:

a. Use four options

Based on the theories, a multiple choice option is said to have good construction if the options have four options. This research analyzes Computer Based English Test for grade VIII. In Indonesia, test for Junior High School students must use four options in multiple choice items. The analysis result shown, all 50 items have four options. The following is an example of these 50 items:

**Figure 4.26 Example of Multiple Choice-Option
of Computer Based English Test
which Use Four Options**

As displayed in figure 4.26 above, it can be seen that this item uses four options (as shown by the red right brace). It can be said that this item fulfills the criteria of validated multiple choice-options number 1 that is "use four options".

b. Have varied location of the right answer according to the number of options.

Based on the theories, a multiple choice option is said to have good construction if the options have varied location of the right answer according to the number of options. The varied location of the correct answer distribution is to avoid too much gap and deviation of the total number between the correct answer in each option. If the items have 50 items and 4 options in each item, the distribution must be around 15 to 16 for maximum number, and around 9 to 10 for minimum number.

The finding shown, each item in Computer Based English Test have 4 options. Nevertheless, the distribution of correct answer in this multiple choice items do not have varied location, since the amount

between the correct answer in each option has much gap and deviation. Also, the location of correct answer almost same in every 2 sequential item number which test the same material.

The distributions of correct answer in these multiple choice items do not have varied location. The following is the distributions of the amount of correct answer in each item-option of Computer Based English Test:

Table 4.10 Correct Answer Distributions
of Multiple Choice Items
in Computer Based English Test

Correct Answer	Number of Items
A	12
B	17
C	16
D	5

As described in table 4.10 above, it can be seen that the amount of maximum number of B option is 17 and minimum number D option is 5. The calculation of this distribution has too much gap between the amount of maximum and minimum number of correct answer in each option.

The location of correct answer almost same in every 2 sequential item number which test the same material. Even, there are 4 items which have unvaried location of the correct answer. For example in items number 22, 23, 24, 25. These 4 items number are located in sequential item number, but have the same location of the correct answer, that is in options C. The following are these 4 items:

SOAL NO `22`

Ukuran font soal : A A

Medan is . . . Than Jambi

- (A) crowded
- (B) crowdeder
- (C) more crowded ←
- (D) most crowded

SOAL NO `23`

Ukuran font soal : A A A
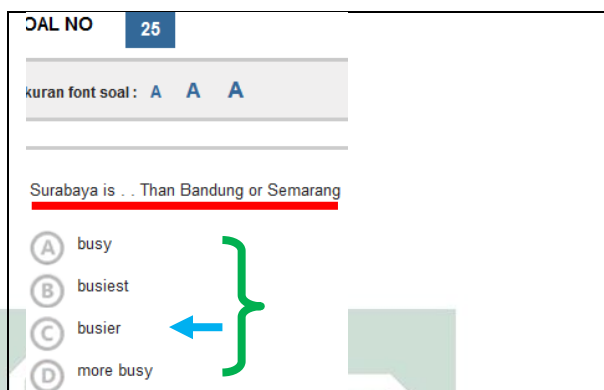
which one is . . .physic or English

- (A) easy
- (B) esier
- (C) easiest ←
- (D) more easy

SOAL NO `24`

Ukuran font soal : A A A

Debu music is . . . Magazine among the teenager in Indonesia

- (A) popular
- (B) more popular
- (C) most popular ←
- (D) famous

**Figure 4.27 Example of Multiple Choice-Option
of Computer Based English Test
which has Unvaried Correct Answer-Location**

As displayed in figure 4.27 above, it can be seen that the options of those 4 items have same location of correct answer (as shown by the blue). All of those 4 items have the key answer located in C option. Even these 4 items test the same material, that is degree of Comparison. The options contain of sentence with different grammar structures which include degree of Comparison, since the options use degree of Comparison words, one of which is "more..." and "...er". Also the stem have Comparison words that is "...than". It means that these 4 items exist to test the same material that is the degree of Comparison. Moreover, the location of putting the correct answer of these 4 items is not random. Students may look for clues that will help them to guess. If the C option is correct answer in the first item, the student will find this pattern and when in doubt choose C option in the next item. It can be said that these 4 multiple choice items of Computer Based English Test do not fulfill the criteria of validated multiple choice-options number 2 that is "Have varied location of the right answer according to the number of options".

c. Options written in numerical order (sequenced).

Based on the theories, a multiple choice option is said to have good construction if the options written in numerical order (sequenced). Options should always be arranged in ascending or descending numerical order to decrease the stress level and minimize taking time of the students. The analysis result shown, from 50 items, there is only 1 item written using numbers and the rest 49 items do not use numbers in the options. So, these 49 items do not need to written in sequential number, since the sequantial number only for options which have number in them.

There is 1 item uses numbers in the options, but the numbers are written in non-sequential. The following is this 1 item with these non-sequential numbers in the options:
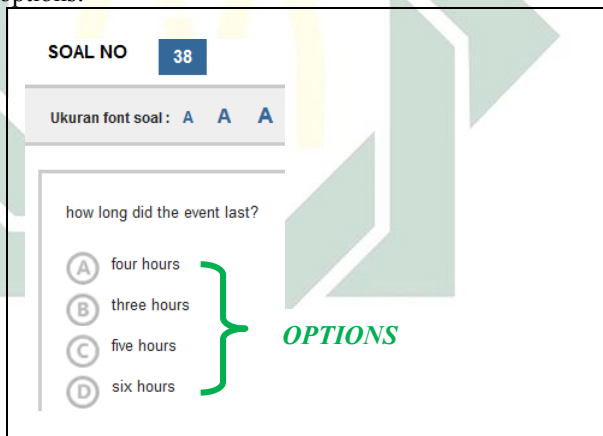

Figure 4.28 Example of Multiple Choice-Option of Computer Based English Test which has Number in the Options

As displayed in figure 4.28 above, it can be seen that this item uses number in the options (as shown by the green right brace). These options are not arranged in ascending or descending numerical order. It can increase

the stress level and take the students' time. It can be said that this item does not fulfill the criteria of validated multiple choice-options number 3 that is "Options written in numerical order (sequenced)".

d. Options written independently; choices should not be overlapping.

Based on the theories, a multiple choice-item is said to have good option-construction if the options used are written independently (not be overlapping). Since, overlapping options give a clue to the students about the correct answer, also if two or more overlapping options are correct, the item may have two or more correct answers. In common, item which has overlapping options written in the form of number. As the example described in the chaper 2.

The researcher found that all 50 items are not in overlapping. Since, from 50 items, only 1 item has number in the options and it does not contain many numbers. The following is this 1 item:
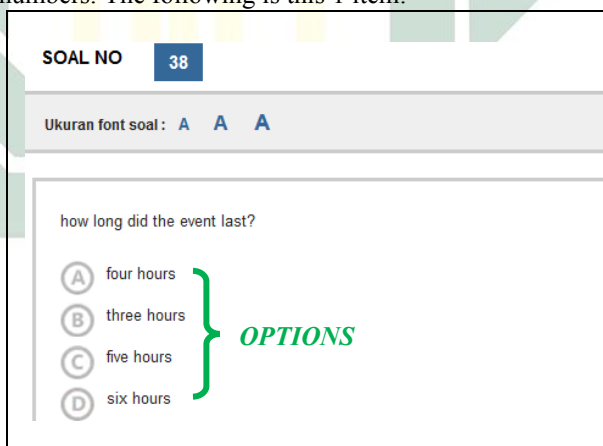
Figure 4.29 Example of Multiple Choice-Option of Computer Based English Test Written with Numerical Options

As displayed in figure 4.29 above, the researcher found that this multiple choice-options use number in the

options (as shown by the green right brace). There is no option in this multiple choice items of Computer Based English Test overlap slightly. It can be said that these all 50 items fulfill the criteria of validated multiple choice-format number 4 that is "Options written independently; choices should not be overlapping".

e. Options homogeneously in content and grammatical structure.

Based on the theories, a multiple choice-item is said to have good option-construction if the options is homogeneous in content and grammatical structure. Since, the use of heterogeneous options in content or grammatical structure may give a clue to the students. The researcher found, there are 29 items have homogeneous options in content and grammatical structure and 21 items have heterogeneous options in content and grammatical structure.

There are 29 items have homogeneous options in content and grammatical structure. From these 29 items, there are 6 items have homogeneous options in content, and 23 items have homogeneous options in grammatical structure. In short, kinds of homogeneous content and grammatical structure covered by these 29 items of Computer Based English Test described in the following table:

Table 4.11 Kinds of Homogeneous Content and Grammatical Structure Covered by Multiple Choice Options of Computer Based English-Test

No.	Kinds of Homogeneous Content and Grammatical Structure
1.	Options written homogeneously in content
2.	Options written homogeneously in content and grammatical structure

As described in table 4.11 above, there are two kinds of homogeneous content and grammatical structure covered by multiple choice options of Computer Based English Test. The following is an example of each kind described in table 4.11:

1) Options written homogeneously in content

There are 6 items have homogeneous options in content. The following is an example of these 6 items:

Figure 4.30 Example of Multiple Choice-Option of Computer Based English Test Written with Options Homogeneously in Content
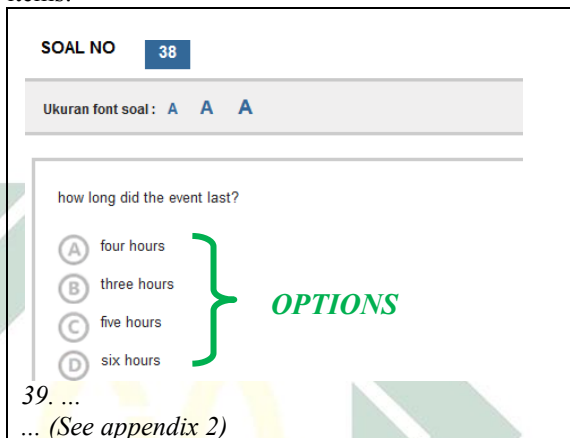
As displayed in figure 4.30 above, it can be seen that this item uses number in the options (as shown by the green right brace). These options are in the same context that is numerical order about "how many hours", since these options use words "four hours", "three hours", "five hours", and "six hours" So, these 4 options written with homogeneously in content. It can be said that this item fulfills the criteria of validated multiple choice-options number 5 that is "Options homogeneously in content".

2) Options written homogeneously in grammatical structure

There are 23 items have homogeneous options in grammatical structure. The following is an example of these 23 items:

**Figure 4.31 Example of Multiple Choice-Option
of Computer Based English Test Written
with Options Homogeneously
in Grammatical Structure**

As displayed in figure 4.31 above, it can be seen that this item uses number in the options (as shown by the green right brace). These options are in the same Grammar structure. The options (as shown by the green right brace) contain of sentence with same element of grammar structures which is quantifier, since these all 4 options use quantifier words which are "there are", "there were", "there is", and "there was". So, these 4 options written with homogeneously in grammatical structure. It can be said that this item fulfills the criteria of validated multiple choice-options number 5 that is "Options homogeneously in Grammatical structure".

There are 21 items do not have homogeneous options in content and grammatical structure. From these 21 items, there are 10 items have heterogeneous options in content, and 11 items have heterogeneous options in grammatical structure. In short, kinds of heterogeneous content and grammatical structure covered by these 21 items are described in the following table:

**Table 4.12 Kinds of Heterogeneous Content
and Grammatical Structure Covered by Multiple
Choice Options of Computer Based English-Test**

No.	Kinds of Heterogeneous Content and Grammatical Structure
1.	Options written heterogeneously in content
2.	Options written heterogeneously in content and grammatical structure

As described in table 4.12 above, there two kinds of heterogeneous content and grammatical structure covered by multiple choice options of Computer Based English Test. The following is an example of each kind described in table 4.12:

1) Options written heterogeneously in content

There are 10 items have heterogeneous options in content. The following is an example of these 10 options:

**Figure 4.32 Example of Multiple Choice-Option
of Computer Based English Test Written with
Options Heterogeneously in Content**

As displayed in figure 4.32 above, it can be seen that these 3 options in the item use Quantifier (as shown by the blue arrow) and 1 option uses

"Article" (as shown by the red arrow). The 3 options which use Quantifier are written in the words, "much", "many", "one" and the 1 option which uses "Article" is written in the word "the". It means that these options are not written with the same content. So, these four options of this item written with heterogeneously in content. It can be said that this item do not fulfill the criteria of validated multiple choice-options number 5 that is "Options homogeneously in content".

2) Options written heterogeneously in grammatical structure

There are 11 items have heterogeneous options in grammatical structure. The following is an example of these 11 items:

Figure 4.33 Example of Multiple Choice-Option of Computer Based English Test Written with Options Heterogeneously in Grammatical Structure

As displayed in figure 4.33 above, it can be seen that these 3 options use Verb Tenses (as shown by the blue arrow) and 1 option uses Gerund (as shown by the red arrow). The 3 options which use Verb Tenses written in the words, "boil", "boils", "have boiled" and the 1 option which uses Gerund

written in the word "boiling". This word "boiling" called to be Gerund, since it has "...ing" after the "Verb", and without "Tobe" before the "Verb-ing". It means that these options are not in the same Grammar structure, since one of the options written with different elemet of Grammar structure. So, this item has heterogeneous options in Grammar structure. It can be said that this item does not fulfill the criteria of validated multiple choice-options number 5 that is "Options homogeneously in Grammatical structure".

f. Have options with the length about the same

Based on the theories, a multiple choice option is said to have good construction if the options have the length about the same. Making the sentence of correct answer longer than the distractors can give clues to the students. The analysis result shown, there are 49 items have options with the length about the same. While, there is 1 item which does not have options with the length about the same.

There are 49 items have options with the length about the same. These 49 items do not consist of too long or too short sentence length in the options. The following is an example of these 49 items:

Figure 4.34 Example of Multiple Choice-Option of Computer Based English Test which has Options with the Length about the Same

As displayed in figure 4.34 above, it can be seen that this item has options (as shown by the green right brace). The length of the correct answer (as shown by the red arrow) does not have striking difference with the length of the distractors (as shown by the blue arrow). So, it may not give a clue to the students in answering the item. It can be said that this item fulfills the criteria of validated multiple choice-options number 6 that is "have options with the length about the same".

There is 1 item does not have the length of options about the same. This 1 item consists of too long and too short sentence length in the options. This 1 item is described in the following:
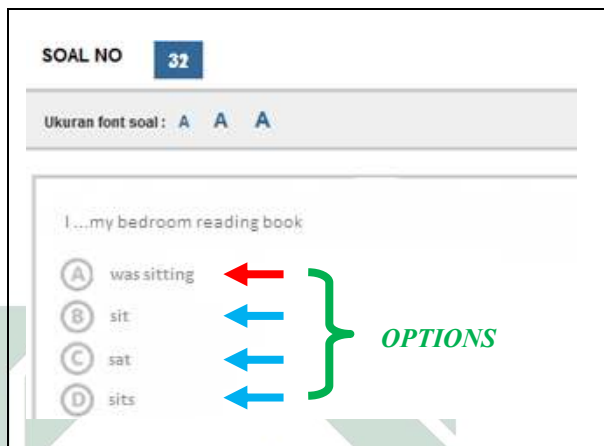
**Figure 4.35 Example of Multiple Choice-Option
of Computer Based English Test which does not have
Options with the Length about the Same**

As displayed in figure 4.35 above, it can be seen that this item has options (as shown by the green right brace). These options are not arranged with the length about the same. The length of the correct answer (as shown by the red arrow) has striking difference with the length of the distractors (as shown by the blue arrow). The correct answer has the longest sentence in the options. So, it may give a clue to the students in answering the items. It can be said that this item does not fulfill the criteria of validated multiple choice-options number 6 that is "have options with the length about the same".

g. Written without using "None of the above".

Based on the theories, a multiple choice-item is said to have good options-construction if the options are written without using "None of the above". Using "none of the above" may force the students to solve the problem rather than choose the correct answer and the students may work backward in using the options. It may take the students' time in aswering the item. The analysis result shown, all the options of 50 items do not use "None of

the above". It can be said that these 50 items fulfill the criteria of validated multiple choice-options number 7 that is "written without using 'None of the above'".

h. Written without using "all of the above" (one right answer).

Based on the theories, a multiple choice-item is said to have good options-construction if the options are written without using "all of the above". Using "all of the above" may give a clue to the students, if student knows that two of the three options listed are correct, that information can clue the students into correctly choosing "all of the above". It seems to draw students into test-taking strategies more than directly testing for knowledge. It means that each multiple choice item should has only one correct answer.

The analysis result shown, all the options of all 50 multiple choice items do not use "all of the above". Nevertheless, from these all 50 items, there are 24 do not have one correct answer. From these 24 items, 2 items do not have the correct answer listed in the options, 2 items have the correct answer which is not suitable with the stated key answer, and 20 items have more than one correct answer. In short, kinds of correct answer covered by these 24 items of Computer Based English Test are described in the following table:

Table 4.13 Kinds of Correct Answer
Covered by Multiple Choice Options
of Computer Based English-Test

No.	Kinds of Correct Answer
1.	Correct answer is not written in the options
2.	Correct answer is not suitable with key answer
3.	Have more than one correct answer

As described in table 4.13 above, there are 3 kinds of correct answer covered in the multiple choice options of Computer Based English Test. The following is an example of each kind of correct answer described in table 4.13:

1) Correct answer is not written in the options

There are 2 items do not have the correct answer listed in the options. The following is an example of these 2 items:

Figure 4.36 Example of Multiple Choice-Option of Computer Based English Test which does not have the Correct Answer Listed in the Options

As displayed in figure 4.36 above, it can be seen that this item has options (as shown by the red right brace). The correct answer of this item is not listed in the options. This item tests Verb Tenses "Past Continuous + Simple Past". The use of past continuous is to discuss an event that occurred at a certain point in the past. Often, the actions described by past simple interfere with the situation described by past continuous. The stem in this item uses incomplete sentence which look for the proper Verb Tense to fill in the gap. In term of this, the stem uses simple past (as underlined by the green line), the pair of the simple past is past continuous (if simple past need pair of Verb Tense). So, the proper Verb Tense (correct answer) is "were laughing" (the form of past continuous). But "were

laughing" is not listed in the options. So, this item does not have correct answer which is listed in the options. It can be said that this item does not fulfill the criteria of validated multiple choice-options number 8 that is "Written without using "all of the above" (one right answer)".

2) Correct answer is not suitable with key answer

There are 2 items have correct answer which is not suitable with the key answer. The following is an example of these 2 items:

Figure 4.37 Example of Multiple Choice-Option of Computer Based English Test which the Correct Answer is not Suitable with the Key Answer
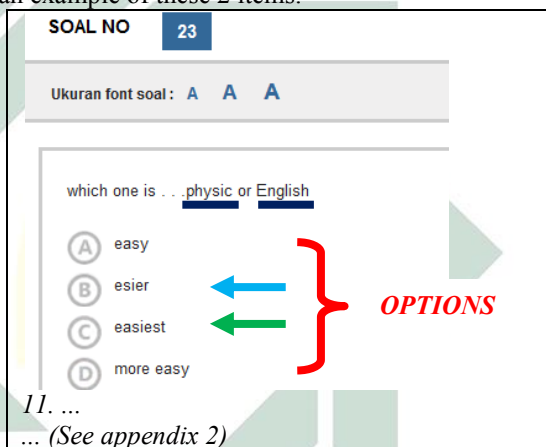
As displayed in figure 4.37 above, it can be seen that this item has options (as shown by the red right brace). The key answer (as shown by the green arrow) is not the suitable correct answer (as shown by the blue arrow). Since, the stem tests Degree of Comparison, which has been marked by two objects compared which are stated in the stem (as underlined by the black line). Degree of Comparsion is used to compare two or more objects that have the same or equivalent nature by using

"more...than" or "....er" The suitable correct answer is "easier". So, this item does not have suitable correct answer. In the other words, the correct answer of this item is not suitable with the key answer. It can be said that this item does not fulfill the criteria of validated multiple choice-options number 8 that is "Written without using "all of the above" (one right answer)".

3) Have more than one correct answer

There are 20 items have more than one correct answers. The following is an example of these 20 items:
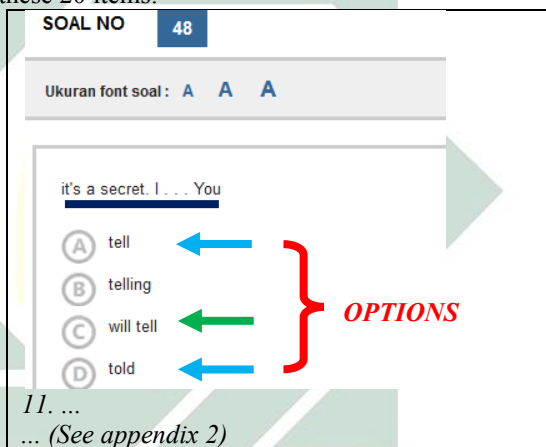
Figure 4.38 Example of Multiple Choice-Option of Computer Based English Test which has more than One Correct Answer

As displayed in figure 4.38 above, it can be seen that this item has options (as shown by the red right brace). The key answer (as shown by the green arrow) is not only the correct answer. The corect answer may be located in the other options "tell" or "told" (as shown by the blue arrow). Since, the stem (as underlined by the black line) does not have time signal to indicate which Verb Tenses is suitbale which the time signal to fill in the gap. Since, the

time signal indicates when the action take place. So, this item has more than one correct answer. It can be said that this item does not fulfill the criteria of validated multiple choice-options number 8 that is "Written without using "all of the above" (one right answer)".

i. Written without using negative words in the options (If a negative word is used, it should be stressed by placing it in bold type, capitalizing it, or underlining it or all of these).

Based on the theories, a multiple choice-item is said to have good options-construction if they are written without using negative words (if a negative word is used, it should be stressed by placing it in bold type, capitalizing it, or underlining it). Using negative word may make the students work backward in using the options. So, it may take the students' time in aswering the item. The analysis result shown, all the options of all 50 multiple choice items do not use negative word. It can be said that these 50 items fulfill the criteria of validated multiple choice-options number 9 that is "written without using negative words in the options".

j. Written each option in grammatically correct when placed in the stem (except the items to test the specific grammar).

Based on the theories, a multiple choice-item is said to have good option-construction if each option written in grammatically correct when placed in the stem (except the items to test the specific grammar). The analysis result shown, there are 45 items which test single language area. From these 45 items, 10 items are written to test vocabulary, 35 items are written to test grammar.

As stated above, each option should be grammatically correct when placed in the stem (except the items to test the specific grammar). There are 10 items test Vocabulary area (do not test specific Grammar area). From these 10 items, only 1 item which one option

is not written in grammatically correct when placed in the stem. The following describes this 1 item:

Figure 4.39 Example of Multiple Choice-Option of Computer Based English Test which the Option is not Written Grammatically Correct when Placed in the Stem

As displayed in figure 4.39 above, it can be seen that this item has options (as shown by the red right brace). These 3 options (as shown by the black arrow) are written with grammatically correct when placed in the stem and 1 option (as shown by green arrow) is not written with grammatically correct when placed in the stem. Since, the stem (as underlined by the blue line) written in incomplete sentence. The students had to fill this incomplete sentence by using the options listed, but there is 1 option written in grammatically incorrect when placed in the stem. Since, this option "he kinds" is the sentence with incorrect grammatical structure. It has to be written "he is kind" then it is grammatically correct when placed in the stem. So, this item consists of option which written in grammatically incorrect when placed in the stem. It can be said that this item does not fulfill the criteria of validated multiple choice-options number 10 that is "Written each option in grammatically correct

when placed in the stem (except the items to test the specific grammar)."

B. Research Discussion

In this chapter, the researcher describes a discussion of the main findings of the result analysis about the construction of multiple choice items. In this section, the researcher describes the research question's answer. It discusses the construction of multiple choice items of Computer Based English Test used in SMP AlKhairiyah Surabaya.

To construct a multiple choice item, there are 3 components that must be fulfilled when constructing a multiple choice item; (1) the format concerns, (2) The writing of stem, (3) and the writing of options.[3] Since those three components which are important in designing a multiple choice items, this discussion categorized into three main findings regarding of those 3 components. Each category has their own criteria to construct good multiple choice items-construction which refers to whether these three categories fulfilled the criteria of validated multiple choice construction or not.

Based on the first category about format concerns, a multiple choice-item is said to have good multiple choice format if the format have met 4 criteria of validated multiple choice items-format, those are; (1)Have stem and options, (2)Have items-format vertically instead of horizontally, (3)Written with correct, punctuation, capitalization, spelling, grammar, and (4) Minimize Reading Time.[4] To construct good format of multiple choice items, the teacher can apply these criteria in writing multiple choice items. In line with that, from the finding of this research, the researcher assumed that from 50 multiple choice items of Computer Based English Test, none of the item has good format, since none of the item fulfills all of those good format concerns criteria. For example, item number 1: between direction and stem sentence of this item structured side by side (not seperated or connect each other). In the other word, between direction and stem of this item become one sentence or the format of this example of item did not format vertically. Also this item has incorrect punctuation. There was no period (.) in the last word of the sentence '5'. 'Do' is the first word

---

[3] Thomas. M. Haladyna, *Developing and Validating Multiple-Choice Test ....*, 97.
[4] Thomas. M. Haladyna, *Developing and Validating Multiple-Choice Test ....*, 98-107.

of the sentence in the stem but there is no punctuation to separate between them as the last word and the first. So, it can combine the direction and the stem-sentence becomes one sentence. This matter may make the students confuse to differentiate between the direction and the stem (the question), also it can make the students misunderstand then make them answered wrongly.

Based on the second category about the writing of stem, a multiple choice-item is said to have good multiple choice stem if the stem have met 5 criteria of validated multiple choice items-stem, those are; (1) Should test single language area (vocabulary and sentence structure should be measured in testing reading skill), (2) Have clear directions, (3) Have brief stem (4) Written without irrelevant information (window dressing), and (5) Written without negative words in the stem (If a negative word is used, it should be stressed by placing it in bold type, capitalizing it, or underlining it, or all of these).[5] To construct good stem of multiple choice items, the teacher can apply these criteria in writing multiple choice items. In line with that, from the finding of this research, the researcher assumed that from 50 multiple choice items of Computer Based English Test, only 1 item has good options, that is number 42, since this item fulfills all of those good options-writing criteria. From the explanation above, it indicates that, the construction of the multiple choice options in item number 42: has single language area, has clear direction, has brief stem, written without irrelevant information, and written without negative words in the stem such as 'NOT' or 'EXCEPT'. So, it can be said that only item number 42 has good stem-writing.

Based on the third category about the writing of options, said to be good multiple choice options if the items-options have met 10 criteria of validated multiple choice items-option, those are; (1) Use four options for Junior High School in Indonesia, (2) Have varied location of the right answer according to the number of options, (3) Options written in numerical order (sequenced), (4) Options written independently; choices should not be overlapping, (5) Options homogeneously in content and grammatical structure, (6) Have options with the same length, (7) Written without using 'None of the above', (8) Written without using 'all of the above', (9) Written

---

[5] Thomas. M. Haladyna, *Developing and Validating Multiple-Choice Test ....*, 107-112.

negative words in the options (If a negative word is used, it should be stressed by placing it in bold type, capitalizing it, or underlining it, or all of these), and (10) Written each option in grammatically correct when placed in the stem (except the items to test the specific grammar).[6] To construct good options of multiple choice items, the teacher can apply these criteria in writing multiple choice options. In line with that, from the finding of this research, the researcher assumed that from 50 multiple choice items of Computer Based English Test, none of the item has good options-writing, since none of the item fulfills all of those good options-writing criteria. For example, item number 1: this item has 3 options which use Quantifier and 1 option uses Article. These 3 options which use Quantifier are written in the words, "much", "many", "one" and the 1 option which uses Article written in the word "the". It means that these options are not written with the same content. So, these four options of this item written with heterogeneously in content.

This research also confirms another study. for example, a study which has done by Iffah Mursyidah Mayangsari. She also found that English Test used is said to have good quality if it was checked by criteria of good test. She investigates about the quality of *UAS English Test of Second Semester 2008/2009 by Teacher-Made Test in SMA 2 Muhammadiyah Sidoarjo* [7]. Here, Iffah analyze the content validity, reliability, difficulty index, and discrimination index of the test content using the criteria of good test as the guidelines. Iffah analyses the Paper Based English test by identifying whether the English test matches with the criteria or not. It similar with what the researcher have done to analyze the English test but in the form of Computer Based English Test. The researcher here used criteria of validated multiple choice items construction to analyze the format, stem, and options of multiple choice items construction. It can be concluded, although the researcher analyses the multiple choice items using different criteria and method also different form of test those are between Paper Based Test and Computer Based Test, criteria here is used as the guidelines to identify whether a test already has good quality or not.

---

[6] Thomas. M. Haladyna, *Developing and Validating Multiple-Choice Test ....*, 112-121.
[7] Iffah Mursyidah Mayangsari, Undergraduate Thesis: "*An Analysis of UAS English Test of Second Semester 2008/2009 by Teacher-Made Test in SMA 2 Muhammadiyah Sidoarjo*" (Sidoarjo : UIN Sunan Ampel Surabaya, 2009).

As the Computer Based English Year End Test is part of the achievement test, it has an important role in measuring the achievement of students' learning outcomes that is the goal of instruction or training,[8] also to summarize what students grasped at the end of a unit of instruction.[9] In line with the important of this test in English asseseement, the quality of this Year-End Test construction made by the teacher to assess students can be prioritized too. If the quality of the multiple choice construction here have not been known whether it includes the multiple choice items that qualify as a good test construction or not. This situation can have an impact on misinterpretation of student achievement outcomes, because the error construction of the test items may make the students more difficult to response (answer) the item, not perform according to the expectations.[10] The criteria of validated multiple choice items construction used here, help the teacher in designing the multiple choice items which is able to know how to construct multiple choice format, stem, and options. Since these category is the main category in building good multiple choice items.[11] That is why both this research and the previous study found that the English test used have good quality or not refers to the criteria.

---

[8] Thomas. M. Haladyna, *Developing and Validating Multiple-Choice Test* ...., 6.

[9] H. Douglas. Brown, *Language Assessment : Principles and Classroom Practice* ...., 6.

[10] J.B. Heaton, *Writing English Language Tests* ...., 37.

[11] Thomas. M. Haladyna, *Developing and Validating Multiple-Choice Test* ...., 97.

CHAPTER V
CONCLUSION AND SUGGESTION

This chapter describes the conclusion of the analysis of multiple choice items of Computer Based English Test used by SMP AlKhairiyah Surabaya to test grade VIII. The section also provides some suggestions for teachers and school party, also future researchers as described in the following:

A. Conclusion

Based on the research findings and discussion of these 50 items-construction including 3 categories in building good multiple choice items, it indicates that: the first category is about the format-construction, none of the 50 multiple choice items have good format-construction, since none of them fulfills all the criteria of validated multiple choice format-construction; The second category is about the stem-construction, only 1 multiple choice item has good options-construction, since this item fulfills all the criteria of validated multiple choice stem-construction; The third category is about the options-construction, none of the 50 multiple choice items have good stem writing-construction, since none of them fulfills all the criteria of validated multiple choice options-construction. As stated above, the following is the construction of each category:

1. In terms of the first category, it indicates that: all 50 multiple choice items have stem and options, and have minimum reading time (it did not take too much students' time). However, all 50 items written with incorrect grammar, punctuation, capitalization, and spelling, and item number 1, 6, 11, 16, 21, 31, 36, 41 are not formatted vertically.

2. In terms of the second category, it indicates that all 50 multiple choice items have brief stem and written without negative word in the stem. However, items number 16, 17, 18, 19, 20 do not test single language area (Vocabulary and Grammar test in one time), items number 1-41, 43-50 do not have clear directions, and items number 2-5, 7-10, 11-15, 17-20, 22-35, 37-40, 43-45, 47-50 written with irrelevant information (window dressing).

3. In terms of the third category, it indicates that all 50 multiple choice items use four options in each items, are not overlapping in writing especially numerical options, are

111

written without using "none of the above" and are written without using negative words in the options. However, the multiple choice items do not have varied location of the correct answer according to the number of options, items number 38 is not written in logical or numerical order (sequenced), the options of the items number 1-5, 11-15, 24, 27-30, 36, 37, 42, 44, 48, 50 do not written homogeneously in content and grammatical structure, item number 32 do not have options with the length about the same, items number 2, 4, 6-8, 10, 11, 13, 16, 17, 18, 19, 20, 23, 30, 31-35, 46, 47, 48, 50 do not have only one correct answer in the options, and the option of item number 44 is not written in grammatically correct when placed in the stem.

B. Suggestion

Based on the analysis result of the construction of 50 multiple choice items of Computer Based English Test, the suggestion of 3 categories here:

1. The first category is suggested for the school, it is better to have an editor to do proofreading to analyze whether the construction of these multiple choice items already have good format or not;

2. The second and the third category are suggested for the teacher to more proritized in analyzing the multiple choice items before using it, and it is important to use the criteria of validated multiple choice items construction to help in designing good multiple choice items-construction;

3. The next is suggested for the item writer (teacher) that all of these 50 items construction should be revised, since each of the item has their own error construction that may impact on misinterpretation of student achievement outcomes, because the error construction of the test items may make the students more difficult to response (answer) the item (not perform according to the expectations).

4. The last is suggested for the future researcher, in this research, the researcher focus on analyzing multiple choice items construction. For future researcher, they can also extend the area of the test form other than multiple choice items, also they can analyze more detail about the construction of Computer Based English Test. Since, this era many

assessment use this method (using application to test the students).
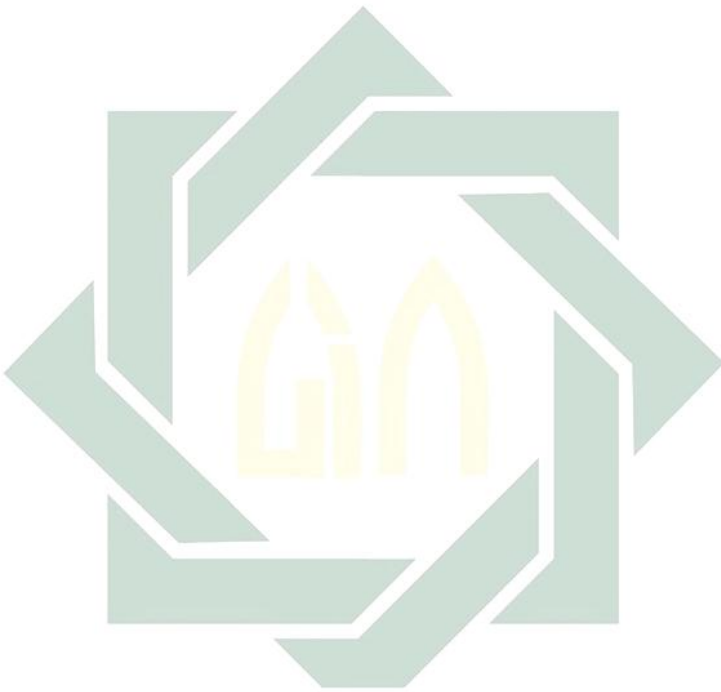
REFERENCES

Bassey, Michael. *Case Study Research in Educational Settings.* Buckingham: Open University Press, 2000.

Boeve, Anja J. *Introducing Computer-Based Testing in High-Stakes Exams in Higher Education: Results of a Field Experiment.* US: US National Library of Medicine National Institutes of Healt, 2015.

Boyce, Jim. *Absolute Beginner's Guide to Microsoft Office*. Indiana: Que Publishing, 2003.

Brown, H. Douglas. *Language Assessment: Principles and Classroom Practice*. San Fransisco: Longman, 2004.

Chin, Christine Hui Li. *The Effect of Computer Based Tests on the Achievement, anxiety, and attitudes of Grade 10 Sciece Students*. Columbia: The University of British Columbia, 1990.

Creswell, J. W. *Research Design: Pendekatan Kualitatif, Kuantitatif, dan Mixed*. Yogyakarta: Pustaka Pelajar, 2012.

DiBattista, David. 2011. "Examination of the Quality of Multiple-choice Items on Classroom Tests". *The Canadian Journal for the Scholarship of Teaching and Learning.* Vol. 2, Iss. 2, Art. 4, December 2011.

Direktorat Pembinaan Sekolah Menengah Pertama Tahun 2017, Kemendikbud.,*Panduan Penilaian oleh Pendidik dan Satuan Pendidikan Sekolah Menengan Pertama*. Jakarta: Kemendikbud, 2017.

Douglas, Dan. *Understanding Language Testing*. Newyork: Routledge Taylor & Fransis Group, 2014.

Ebel, R.L. *Essentials of Educational Measurement, Fifth-Edition.* New Delhi: Prentice Hall of India, 1991.

114

Field, Marion. *Improve your Written English, Fifth-Edition.* Begbroke: How to Content, 2009.

Flesch, Rudolf. *The Art of Readable Writing*. Evanston: Harper & Row Publisher, 1949.

Gray, Paul S. *The Research Imagination: An Introduction to Qualitative and Quantitative Methods*. New York: Cambridge University, 2007.

Grondlund. Norman E, *Assessment of Students Achievement, Seventh-Edition*. Boston: Pearson Education, 2003.

Gronlund, Norman E. *Constructing Achievement Tests, Second-Edition.* London: Prentice-Hall, Inc., Englewood Cliffs, 1977.

Haladyna, Thomas. M. *Developing and Validating Multiple-Choice Test Items, Third-Edition*. New Jersey: Lawrence Erlbaum Associates, 2004.

Harmer, Jeremy. *The Practice of Language Teaching, Third-Edition*. London: Longman Group, 2007.

Heaton, J.B. *Writing English Language Tests*. New York: Longman, 2000.

Kubiszyn, Tom. *Educational Testing and Measurement*, *Sevent-Edition*. Newyork: John Wiley & Sons, Inc, 2003.

LaRocque, Paula. *The Book on Writing: The Ultimate Guide to Writing Well.* Portland: Marion Street Press, 2003.

Mayangsari, Iffah Mursyidah., Undergraduate Thesis: "*An Analysis of UAS English Test of Second Semester 2008/2009 by Teacher-Made Test in SMA 2 Muhammadiyah Sidoarjo"*. Sidoarjo: UIN Sunan Ampel Surabaya, 2009.

McCaskill, Mary K. *Grammar, Punctuation, and Capitalization: A Handbook for Technical Writers and Editors*. Hampton: Langley Research Center, 1998.

Miller, M. David, - Roberet L. Linn. *Measurement and Assessment in Teaching, Tenth-Edition*. New Jersey: Pearson Education, Inc.,Upper Saddle River, 2009.

Mehrens, William. A, - Irvin Lehmann. J. *Measurement and Evaluation in Education and Psychology, Fourth-Edition*. Belmont: Wadsworth, Thomson Learning, 1991.

Merina, Hana Dyah., Undergraduate Thesis: "*Analyzing The Multiple-Choice Items of The Reading Test in Senior High School National Examination*". Yogyakarta: Sanata Dharma University, 2009.

Mukherjee, Poulomi. 2015. "Analysis of MCQs Item & Test Statistics from an assessment in a medical college of Kolkata, West Bengal". *IOSR Journal of Dental and Medical Sciences (IOSR-JDMS)*. Vol. 14, No. 12, Dec. 2015. 47-52.

Namdeo, Surya Kumar. 2016. "Item analysis of multiple choice questions from an assessment of medical students in Bhubaneswar, India". *International Journal of Research in Medical Sciences*. Vol. 4, No. 5, May 2016. 1716-1719.

Nurliyanto, Dian, Undergraduate Thesis: "*Test Item Analysis of The Final Examination on Economics Subject Grade XII IPS SMA Negeri Banyumas Academic Year 2014/2015*". Yogyakarta: Yogyakarta State University, 2015.

Oosterhof, Albert. *Developing and Using Classroom Assessment, Third-Edition*. New Jersey: Merrill/Prentice Hall, 2003.

Osterlind, Steven J. *Constructing Test Items: Multiple-Choice, Constructed-Response, Performance, and Other Formats, Second Edition*. Moscow: Kluwer Academic Publisher, 2002.

Permendikbud, Kemendikbud.,*Standar Penilaian Pendidikan,* BAB V Pasal 6: Bentuk Penilaian. Jakarta: Kemendikbud, 2016.

Pusat Penilaian Pendidikan, Kemendikbud.,*Panduan Penulisan Soal Pilihan Ganda*. Jakarta: Kemendikbud, 2017.

Shidiq, Dja'far - Halfian Lubis. "The Technique of Multiple Choice Answer Key Randomization" *Dunia adalah Ruang Belajar Sesungguhnya*, (https://rofiquez.wordpress.com, accessed on January 01, 2019)

Smith, Brady. *Proofreading, Revising, & Editing Skills Success in 20 Minutes a Day.* Newyork: Learning Express, LLC, 2003.

Thayn, Kim Scott., Doctor of Philosophy Dissertation: "*An Evaluation of Multiple Choice Test Questions Deliberately Designed to Include Multiple Correct Answer*". Provo: Brigham Young University, 2010.

The Result of Interview with English Teacher-Mrs. Nyimas. Surabaya: SMP AlKhairiyah Surabaya, 2018.

Wyatt, Claire, Smith, and Joy Cumming. *Educational Assessment in the 21st Century: Connecting Theory and Practice*. London: Springer, 2009.

M. Downing, Steven, - Michael C. Rodriguez. *A Review of Multiple Choice Item-Writing Guidelines for Classroom Assessment.* New Jersey: Lawrence Erlbaum Associates, 2004.