

# ESP IN PRACTICE

*by* Kun Aniroh

---

**Submission date:** 28-Jun-2018 07:59AM (UTC+0700)

**Submission ID:** 979038036

**File name:** ESP\_IN\_PRACTICE.docx (297.85K)

**Word count:** 33759

**Character count:** 185615

**ESP IN PRACTICE :DESIGNING READING  
TEST FOR HOTEL STUDENTS**

**Kun Aniroh Muhrofi-Gunadi**

### Acknowledgement (letaknya di depan).

<sup>33</sup> I would like to convey my sincere gratitude to Allah Subhanahu Wa Taala, The Most Gracious, The Most Merciful, who have given me opportunity to accomplish all the duties during<sup>47</sup> my study.

I would like to express my appreciation to a number of a number of people who helped me to write the book. This book is an assignment<sup>71</sup> of English Language Testing as one of the courses of English Education Program. In this opportunity I would like to convey my gratitude to Prof. Dr. Soenardi Djiwandono, MA.,Ph.D who assigned me to make the Language Testing project at the Graduate English Education Program of Universitas Negeri Malang. Without the assignment the book does not come into existence. I am also thankful to the Head of English Study Program of State University of Maulana Malik Ibrahim Malang (UIN) who gave permission to do the try out.

I also convey my gratitude to the students of Hotel Department of Diploma III Tourism Program for the test that I conducted and the students of English Department of English Department of UIN for the try out. Another gratitude goes to Pak Masduki, Pak Indawan, Bu Sinta, Pak Agus, my colleagues at the Graduate English Education Program who always support me and one another to conduct the project, it was really sweet memories in struggling to do the assignment which sometimes made us having the feeling in between of hope and being in despair.

Last but not least in doing the assignment I always accompanied by my lovely kids Pipiet, Neno and Fiki thank you for their support and silent prayer, my husband 'Papie Gunadi' for the guidance and proofreading, Mbak Surati, Mbak Yah who helped me in the 'family domestic and foreign affairs' and Pak Tris the driver, who took me to the work and the campus.

This work is solely my own responsibility for the mistakes, errors and inadequacies. Criticism and feedback will be most welcome.

Malang, June 2018

Kun Aniroh Muhrofi-Gunadi

## FOREWORD (depan 2)

Quality assurance in education in its most genuine sense should take place nowhere but in its very habitat: the classroom practice. Whatever innovative educational reform programs introduced to classroom practitioners seems to have no substantial effect on attempts made to change classroom culture for quality improvement in education if classroom practitioners were not aware of what and for what they are to perform professionally in the classroom. As the critical agent of culture change in the classroom, teachers need to be professionally and adequately equipped with up-to-date knowledge and skills in assessing whatever products resulted from their instruction. One is knowledge and skills in testing and interpreting the learning outcomes.

It is under such a vision that the present final project is accomplished. The report is written in 5 (five) chapters. Chapter I introduces the context for accomplishing the final project. Chapter II lays both theoretical and conceptual foundations of aspects related to reading and language testing. Chapter III describes the procedures undertaken to carry out the test validation. Chapter IV reports all accounts pertaining to validation. Chapter V concludes the conduct of validation and recommends further actions needing to be taken in conjunction with the findings of the final project.

A lot is learned from accomplishing this final assignment. We can reflect vividly what and how we, classroom practitioners like an amateur, have tested and evaluated our students with our own-made 'invalidated' tests so far. Quality assurance in education should accordingly emanate from quality classroom practices: from quality tests made by reflective teachers. Not others. So, start small, Begin now, Assess and Evaluate!

Malang, 24 June 2018

The Writer,

Kun Aniroh Muhrofi-Gunadi



## **KETERANGAN SAMPUL BELAKANG BUKU**

This book is designed for the advanced students of English Department, English instructors and those who are interested in test of English for specific purpose. It is designed based on the principles of language testing with the content of hotel and tourism matters. In designing summative test for the needs of students' quizz, middle or final test, teacher made test is usually lack of implementing the test principles. This does not mean that in the teacher made test there should be an 'ideal made test' but rather to learn of how to make 'ideal test' and can be practiced for the broader needs such as students' entrance test, staffs' selection and recruitment test and other crucial need test in an institution.

Theoretical overview of basic concept of reading, taxonomy of reading, assessment, format, requirement of a good test is provided which enrichs the reader in comprehending the theory of reading comprehension and reading test. The procedure for test development deals with purpose of testing reading, test content specification, text selection, scoring grid, test item moderation and test item revision.

Try out and result which discuss analysis of try out test, item analysis, test validity and practicality are also presented. To provide the readers with the statistical procedure, list of figures of the test procedure and .... are found at the appendices of the book.

**ABOUT THE WRITER ( ada foto) insha Allah saya kirim lewat WA., letakkan di halaman akhir sebelum sampul belakang**

Dr.Kun Aniroh Muhrofi-Gunadi, SST.Par.,MM.,MP.d is a lecturer at Diploma Tourism Program, Universitas Merdeka Malang, Indonesia. Her education backgrounds are English, human resources and tourism management and got the doctorate degree of English education of Universitas Negeri Malang. Her teaching and research interests focus on TEFL, ESP, materials development and tourism destination management. She has written a number of articles in journals and proceedings and had written two books ESP Materials Development: Theory and Practice and English for Hotel Management.

In 2004, she had the opportunity to attend fellowship program in Belgium and Germany, in 2006 attended Ecotourism short course in Japan, and in 2010 was a visiting lecturer at English Education of Kyungpook National University and at English and International Tourism Department of Gyeongju University, Korea. In 2012 she was a visiting lecturer at the Faculty of Business and Tourism at Universiti Malaysia Kelantan and in 2016 she was invited by the Consulate General of Indonesia in Perth, Australia to share about social community development.

She was the volunteer of The Indonesian Foundation for Intercultural Learning (2002-2017) and was a member of board of advisor of AIESEC (2008-2017) of Brawijaya University, Malang.

She has been invited to be a resource person in talkshow, conference and seminars by some organizations discussing English teaching, tourism issues, research and social activities. She is the founder of PINOKI, a foundation of study club, English and Quranic study for kids, boarding house, syaria cooperative for villagers and ecotourism in the village of Klandungan, Landungsari, Malang.

## TABLE OF CONTENTS

<b>CHAPTER I</b>	<b>INTRODUCTION .....</b>	<b>1</b>
1.1	Background .....	1
1.2	Rationale for Reading Skill Instruction .....	2
1.3	Challenges and Constraints .....	3
1.4	Objectives of the Final Project.....	3
<b>CHAPTER II</b>	<b>THEORETICAL OVERVIEW.....</b>	<b>5</b>
2.1	Basic Concepts of Reading, Reading Processes and Comprehension .....	5
2.2	Factors Influential in Reading Comprehension .....	5
2.3	Taxonomy of Reading.....	7
2.4	Assessment of Reading Skills in ESP .....	8
2.5	Formats of Tests of Reading Skills .....	9
2.6	Requirements of a Good Test.....	10
<b>CHAPTER III</b>	<b>PROCEDURES FOR TEST DEVELOPMENT .....</b>	<b>13</b>
3.1	Purpose of Testing Reading Based on the Course Description .....	13
3.2	Subjects of the Course.....	13
3.3	Test Content Specification .....	14
3.4	Selection of the Test Format .....	16
3.5	Text Selection .....	17
3.6	Item Writing and Item Assembling.....	19
3.7	Writing the Key Answers and Test Scoring Guide.....	22
3.8	Grading and Grade Interpretation .....	25
3.8.1	Grading.....	25
3.8.2	Grade Interpretation .....	26
3.9	Test Item Moderation.....	27
3.9.1	Moderation Processes.....	28
3.9.2	Moderators' Feedback and Its Analysis.....	30

3.10	Test Item Revision .....	35
3.11	Informal Tryout.....	38
3.11.1	Administration of the Informal Tryout.....	38
3.11.2	Analysis of Informal Tryout Results.....	38
3.11.3	Follow-up Steps.....	41
<b>CHAPTER IV READING TEST TRYOUT AND RESULTS.....</b>		<b>42</b>
4.1	Reading Test Administration .....	42
4.1.1	Place and Time of Administration .....	42
4.1.2	Test Takers .....	42
4.2	Analysis of Tryout Test Results.....	43
4.2.1	Description of Students' Achievement.....	43
4.2.2	Item Analysis.....	47
4.2.3	Evidence of Test Reliability .....	49
4.2.4	Test Validity .....	51
4.2.5	Evidence of Practicality .....	64
<b>CHAPTER V CONCLUSION AND SUGGESTION.....</b>		<b>66</b>
5.1	Test Development in Retrospect: Conclusions .....	66
5.2	Suggestion.....	67
<b>REFERENCES .....</b>		<b>68</b>
<b>APPENDICES .....</b>		<b>70</b>

## LIST OF TABLES AND FIGURES

Table 3.1: Test Content Specification.....	16
Table 3.2: Approximate Requirements of Texts.....	19
Table 3.3: Final Item Placement and Their Abilities Referred to.....	21
Table 3.4: The Scoring Scheme.....	24
Table 3.5: Score Weighting Calculation.....	25
Table 3.6: Grade Conversion Table.....	27
Table 3.7: Summary on Moderator's Feedback.....	30
Table 3.8: Table Analyzing Moderator's Feedback.....	31
Table 3.9: Summary of Statistics for Moderators' Consistency.....	33
Table 3.10: Agreement of the Moderator's Responses.....	34
Table 3.11: Revisions on Instructions.....	35
Table 3.12: Revisions on Item Wording.....	36
Table 3.13: Post-Informal Trial: Students' Feedback.....	40
Table 4.1: Summary of Statistical Descriptions of Students' Achievement.....	44
Table 4.2: Frequency of Students' Grade Occurrences.....	45
Table 4.3: Students' Evaluation on the Test Tryout.....	46
Table 4.4: Index and Category of Item Difficulty.....	48
Table 4.5: Index and Category of Item Discriminating Power.....	48
Table 4.6: Summary of Results of Item Analysis.....	49
Table 4.7: Summary of Results of Reliability Analysis.....	50
Figure 4.1: The Structure of Reading Ability Conceptualized.....	54
Table 4.8: Unrotated Component Matrix.....	57
Table 4.9: Rotated Component Matrix.....	58
Table 4.10: Summary of Concurrent Validity Analysis.....	60
Table 4.11: Summary of Results of Item Analysis.....	62

## LIST OF APPENDICES

Appendix 1: Reading Text Pool for Initial Text Selection .....	70
Appendix 2: Test Draft for Moderation .....	73
Appendix 3: Moderators' Evaluation Feedback Sheet.....	79
Appendix 4: Test Paper for Informal Trial/Pilot Testing.....	81
Appendix 5: Students' Evaluation Feedback Sheet.....	87
Appendix 6: Respondents' Responses to All the Test Items during Informal Trial.....	88
Appendix 7: Final Version of the Test.....	89
Appendix 8: Answer Sheet of the Final Test Form .....	95
Appendix 9: Final Scoring Guide .....	98
Appendix 10: Raw Score Recording Sheet.....	101
Appendix 11: Students' Data Recording Table .....	102
Appendix 12: Item Analysis Results.....	103
Appendix 13: Complete Manual Computation .....	104
Appendix 14: Results of Descriptive Analysis .....	105
Appendix 15: Result of Factor Analysis.....	108
Appendix 16: Sample Student's Work on Piloting Stage .....	116
Appendix 17: Sample Student's Work on Test Tryout.....	118



## **CHAPTER I**

### **INTRODUCTION**

#### **1.1 Background**

The teaching of Reading Comprehension at Diploma III in Tourism Program is still posed with several technical and non technical challenges. This is partly, it should be genuinely admitted, due to lack of quality resources such as human resources as well as instructional materials to be taught to students either at lower level and higher level. At both levels, for instance the reading materials for students of tourism program are limited to the old materials and some selected materials prepared by the teachers without conducting careful need analysis. Consequently the teaching learning process of reading comprehension does not give satisfactory results as shown by their achievement (Official Documents on Students' Achievement Results, 2004). An attempt has been made so far by the English teachers at Diploma III in Tourism Program and the subject specialists to bridge this by compromising some topics to be taught and by teaching English while students have practice on some subject skills, but again this does not give satisfactory output yet.

ESP reading skills actually place a strategic position in tourism industry, in particular for those at the managerial levels. This means that students who have good reading comprehension skill will have good position in their future career. Reading comprehension is needed at all departments in hotel and tours and travel but high level of reading is mostly needed at the 'back office' which needs careful analysis and critical thinking on the part of the students. Back office is a department such as front office, marketing, human resources at the hotel, and marketing, ticketing and tour department at tours and travel. As a consequence, those who want to have a good career at tourism industry should have high level of reading.

Since the position of Reading Comprehension is crucial in tourism industry, it is hoped that not only teaching learning processes which should be taken into consideration, but also the testing. A good test can have a good impact on teaching learning process, materials and the objective of testing (Bachman, 1990:283; Hughes, 2003:1-2; Brown, 2004:26). Brown (2004:26) even suggests that teaching learning process starts from testing. By so doing the direction of teaching is clear, and the students know what they should do at the end of the lesson or the semester. A good test can also be the source of learning for the students. From the result of the test they can improve their mistakes they have made, so they will not make the same mistakes.



A formative of ESP Reading Comprehension test here is proposed for the sake of practicality. First, it is because the test is conducted while an instructional program is going on. The item coverage is limited to what has already been given in the program. This is done to see whether the program is running in line with its plans. Improvement may be set up if the program goes beyond the plans. Second, the ESP Reading Comprehension Test is specific, usually contrasted with a language test in general the purpose of which is broadly defined. A language test in Language for specific Purpose Testing (LSP) such as the ESP Reading Comprehension Test has the authenticity of task and the interaction between language knowledge and specific purpose content knowledge (Douglas, 2000:2).

## 1.2 Rationale for Reading Skill Instruction

The kind of English taught to students at Diploma III Tourism Program so far has been English for Specific Purposes, particularly the kind of English that is concerned with tourism industries. The reason for equipping the students with ESP of tourism industries is that by the end of their study the students are expected to have been functionally capable in using English for their prospective professions in tourism industries. For the purpose, English is allocated for 16 credits, which is distributed into courses. These courses are Communicative English and Professional English. These courses have objectives different emphases. The former focuses on the mastery of listening and speaking skills whereas the latter puts a more emphasis on reading and writing skills. To evaluate the students' achievement in Communicative English(Speaking) and Professional English(Reading), teacher-made versions of test have been employed. Normally, the tests are developed by the lecturers responsible for the corresponding courses. The test items are developed based on instructional materials outlined in the syllabus; no test content specification has been developed so far as the basis of writing test items. Similarly, there has been no initial trial to be applied to examine empirically the quality of the items and the test in terms of item analyses, reliability and validity aspects of the tests. Thus, the tests are basically intuitions of the lecturers. As such, the quality of the tests is open to questions and so is the interpretation of the scores. Considering the important role of tests in measuring the students' genuine English skills, there is a need to examine more closely the quality of the tests as used in Diploma III Tourism Program.

### **1.3 Challenges and Constraints**

The challenges and constraints meant in this context refer to external and internal factors. In external terms, the Language for Specific Purpose (henceforth LSP) Test is relatively new. LSP Test was introduced in 1975 by the British Medical Council for the purpose of evaluating the professional and language abilities of physicians trained outside UK applying for temporary registration to practice medicine in Britain (Douglas, 2000:3). Therefore, if the test has not been well established, it is quite acceptable. An LSP test covers many kinds of field in which each field has its own characteristics in language and content. For this reason it is not easy to standardize an LSP Test. An LSP Test in Tourism which is commonly used is not standardized in the sense that every industry has its own standard when the industry wants to recruit new staffs. This becomes a challenge as well as a constraint for tourism colleges. This means that English teachers are demanded to master many kinds of genre, language and content in Tourism.

Internal challenges and constraints are also observed. The test of reading comprehension is one of the tests used in Diploma III Tourism Program. The reading skills are considered more strategic particularly in equipping the students in Professional English because these students are projected to be able to function more as supervisors or middle managers where reading skills are required. The internal challenges and constraints seem to be highly technical. There has been absence of assuring test quality.

Considering the absence of systematic attempts in the test development in Diploma III Tourism in general as well as the important role of reading tests in Professional English, there is a strong call for a more systematic endeavor to develop such reading tests. This includes the process from establishing the test content specification of the intended reading tests, the writing of test items based on the test content specification, the trial or validation of the test to examine the quality of the items that covers discrimination power, level of difficulty as well as effectiveness of distracters, the reliability and the validity of the reading test.

### **1.4 Objectives of the Final Project**

The reading comprehension test to be developed is intended for the fifth semester of Hotel Department students, Diploma III in Tourism Study Program, Merdeka University Malang. As has been described in the previous section, simply no quality assurance has been established and exercised in all lecturer-made tests in Diploma III in Tourism Study Program, Merdeka University Malang. As a result, these tests are questionable in terms of their

characteristics such as level of difficulty, discrimination, reliability, and validity. Considering that tests conceptually are to have sequential effects (Brown, 2004:26), be it on lecturers' teaching and students' learning, letting the tests be administered lacking of validation seems to be not desirable if quality learning output is expected. This disadvantaging situation certainly needs to be avoided at all points. Therefore, a fundamental change in practice needs to be initiated immediately.

The aim of the final project is mainly two folds: to develop an LSP Test of reading comprehension using correct principles of test, and then to conduct validation on the reading test to be developed. The test quality to be examined include item characteristics, test reliability, and test validity with their specific dimensions such as item level difficulty, item discrimination, <sup>82</sup> internal consistency, face validity, content validity, concurrent validity, and construct validity as well as test practicality.

## CHAPTER II

### THEORETICAL OVERVIEW

#### 2.1 Basic Concepts of Reading, Reading Processes and Comprehension

Alderson (2000:3-4) presents a distinction between the process at reading and the result of the process, the product. The process is what we mean by 'reading proper': the interaction between the reader and the text. During that process, presumably, many things are happening. The reader is also thinking about what he is reading: what it means to him, how it relates to other things he has read, to things he knows, to what he expects to come next in text like this.

Many different things can be going on when a reader reads; the process is likely to be dynamic, variable, and different for the same reader on the same text at a different time or with different purpose in reading. The process will be different for different readers on different texts at different time or with a different purpose of reading.

There are two different approaches that may be taken by the readers. One is the bottom-up approach and the other one is the top-down approach. Bottom-up approaches are the serial models, where the reader begins with printed word, recognizes graphic stimuli, decodes them to sound, recognizes words and decodes meanings. Top-down approaches emphasize the importance of the schemata and the readers' contribution over the incoming text. Schemata or schema-theoretic models are based upon the theory which accounts for the acquisition of knowledge and interpretation of a upon text through the activation of schemata. In this view readers activate what they consider to be relevant existing schemata and map incoming information onto them. To the extent that these schemata are relevant, reading is successful (Alderson, 2000:17)

#### 2.2 Factors Influential in Reading Comprehension

Conceptually, there are several important factors affecting reading comprehension (Bachman, 1990; Grabe, 1991). According to Grabe (1991), there are six factors: automatic recognition skills, vocabulary and structural knowledge, formal discourse structure knowledge, content/world background knowledge, synthesis and evaluation skills/strategies and meta cognitive knowledge and skills monitoring. In a more operational term, the function of these factors plays an essential role that enable reading comprehension to happen. Reading



comprehension will not take place if there exists insufficiency or incompetence in one or more of these factors. Put simply, if comprehension is understood as understanding text contents which requires the activation of those factors, then, the dimension of reading comprehension in this context is concerned with the understanding of text contents. Text contents in the present test refer particularly to these texts of the intermediate to upper intermediate levels of a particular genre (Celce-Murcia, Dornyei, and Thurrell, 1995); that is, the ones related to tourism industries, particularly hotel-management matters.

There are several functional competencies that a efficient reader needs to master. Munby (1978) called these skills as micro reading skills. Conceptualizations of functional reading competencies as constructs other than that proposed by Munby are also used. The ones in standardized proficiency tests like the TOEFL (Gear and Robert, 1996) as well as the IELTS (Alderson, 2000:131) are obvious. The reading course in the present report is designed to include some of the functional competencies proposed by Munby (1978), the TOEFL (Gear and Robert, 1996) and the IELTS (Alderson, 2000:131). The rationale is that these functional competencies seem to be so general and essential that a good reader needs to possess regardless their purpose of reading. These functional skills include locating factual information, determining propositional inferences of a text, evaluating expressions in a text, summarizing and inferring underlying information in a text and understanding information in a text through visual presentations. Accordingly, these "micro skills" outlined above are considered to be reading skills functional for the purpose of measurement in the present test.

Conceptualized functional reading skills are essentially abstract. That is, they cannot be directly measured. For the purpose of measurement, these skills need to be operationally defined in terms of indicators that reflect the essence of these functional reading skills. Wiener and Bazerman (1988) and Anderson, Durston and Poole (1969) provide a clear framework on how these functional skills may be defined. A critical analysis on their concept yields indicators for each of these skills as follows: Locating factual information includes finding main factual information: *who*, *what*, *where*, and/or *when*; and finding factual explanatory: *why* and *how*; determining propositional inferences of a text gives in finding propositional informative inferences, finding propositional exploratory inferences; evaluating expressions in a text covers identifying values of expressions in the forms of facts, opinions, evidence, definition, or implication; summarizing and inferring underlying information in a text provide indicators as drawing a logical conclusion, predicting outcomes, and/or generalizing; finally, understanding information in a text through visual presentations has as

its performance indicator 'presenting information contained in a text through visual presentations like a picture, a chart, or a curve.'

### 2.3 Taxonomy of Reading

As quoted by Champeau et. al (1997:31-32) there are some taxonomies which are applicable for reading comprehension. They explained it from Bloom's taxonomy which was applied to the areas of instruction and evaluation. Since the taxonomy was very influential, in its development it is then applied in reading comprehension. Here are the some taxonomies of reading which can be taken into consideration to develop reading comprehension test.

The first taxonomy, and probably the best known was published by Bloom et.al, in 1956. The main purpose of the taxonomy was to classify educational objectives, but it was later also applied to the areas of instruction and evaluation. It is divided into three large areas or domains, the cognitive domain, the affective domain and the psychomotor domain.

In 1978, Herber tried to relate Bloom's categories to the three levels of reading comprehension: literal comprehension, interpretive comprehension, and applied comprehension. Literal questions require the reader to recall or recognize information explicitly presented in the reading material. Interpretive questions ask for a paraphrase, explanation, inference, conclusion, or summary. Applied questions utilize the reader's background knowledge and lead them to evaluate, elaborate, predict or solve problems based on implicit information in the text.

In 1978, Pearson and Johnson presented taxonomy of word comprehension tasks with nine levels of taxonomy of propositional comprehension tasks. Their questions of taxonomy consist of three levels: textually explicit questions, textually implicit questions, and scriptically implicit questions. The definitions of these categories correspond roughly to those of Herber (1978) and Gray (1960) has called to "read the lines, read between the lines, and read beyond the lines."

Besides, there are other kinds of reading taxonomy. Barrett's taxonomy (1968) also refers to questions related to reading comprehension and is far more detailed. There are four main categories: literal recognition or recall, inferences, evaluation, and appreciation. Elijah and Legenza (1975) present a taxonomy based largely on Barrett's and Sander's (1966) publications, They also describe four main levels of comprehension (literal, interpretive, reaction, and application) with numerous subcategories. Irwin's taxonomy (1986) best reflects the interactive theory of reading comprehensions. He separates questions at the level of

micro-information explicitly presented in the reading material (concerning word meaning or syntactic relationships) from the questions at the level of macro-information (main ideas summaries). These taxonomies have different characteristics. This means that a certain type of taxonomy will be applicable for a certain type of a certain type of format

## 2.4 Assessment of Reading Skills in ESP

Over the years since the beginnings, specific purpose for language testing has been criticized on a number of grounds; specific purpose language proficiency is really just general purpose language proficiency with technical vocabulary thrown in; if we test general language knowledge, specific uses will take care of themselves. However since ESP test has been construed as those involving language for academic purposes and occupational or professional purposes, the specific purpose language tests are indeed necessary, reliable, valid and theoretically well-motivated (Douglas 2000, 1-2).

There are some reasons why specific purpose language tests are necessary : researchers are pretty much in agreement that language performances vary with both context and test task, and therefore our interpretations of a test taker's language ability must vary from performance to performance; technical language that is used in any academic, professional or vocational field including cooking, law, physics, chemistry, air-traffic control, scuba diving, stamp collecting, or language teaching has specific characteristics that people who work in that field must control (Douglas 2000,6-7).

In relation to reading skills in ESP Alderson (1995, in Douglas 2000:168) states that an alternative model involving stages and construction and evaluation covers: identifying test purpose; developing test specifications; guidelines for training of item/task writers and moderation of their products; pre-testing analysis of revision test; training examiners and administrators; monitoring examiner reliability; reporting scores and setting pass marks; test validation; post-test reports and developing and improving tests.

For the aspects of specifications ESP test should contain statements covering all or most of the following points: test purpose, the learner taking the test (age, sex, level of language proficiency, first language, cultural background, country of origin, educational level and nature of educational reason for taking test, likely personal and professional interests and level of background knowledge, etc), test level (in terms of test-taker ability), test construct, description of suitable language course or textbook, number of sections to the test, time for each section, target language situation(s), text types, text length, text complexity/difficulty,

language skills to be tested, language elements (structures/lexis/notions/functions), task types, number and weight items, test methods, rubrics, examples; explicit assessment criteria, criteria for scoring description of typical performance at each level, description of what candidates at each level can do in the real world (Palmer and Bachman 1996 in Douglas 2000:169)

## 2.5 Formats of Tests of Reading Skills

There are many types of test in reading, and in this discussion only some will be addressed to focus on short answer or restricted response. Heaton (1975:107-134) divides reading into initial stages of reading, intermediate and advanced stages of reading, and open-ended and miscellaneous items. Initial stages of reading consist of word matching, pictures and sentence matching. Intermediate and advanced stages of reading consist of true/false reading tests, multiple-choice items: short texts, multiple choice items: longer texts, completion items, rearrangement items, cloze procedure, and open-ended and miscellaneous items.

Alderson (2000:207-227) presents cloze test and gap-filling tests, multiple-choice techniques, matching techniques, ordering tasks, editing tasks, alternative integrated approach, short-answer tests, free-recall test, the free-recall test, the summary test, information-transfer techniques, real-life methods, and informal methods.

Short answer test simply takes a question which requires a brief response in a few words. The justification of this technique is that it is possible to interpret students' response to see if they have really understood. The objectivity of scoring depends upon the completeness of the answer key and the possibility of students' responding with answer or wordings which were not anticipated. Short-answer questions are not easy to construct. The question must be worded in such a way that all possible answers are foreseeable. Otherwise the marker will be presented with a wide responses which she will have to judge as to whether they demonstrate the understanding or not.

Nitko (1996:160-176) mentions essay assessment tasks. These consist of restricted response, multiple-choice interpretive exercises and extended response varieties. Restricted response items restrict or limit what you will permit the student answer. The way you phrase a restricted response task requires a student to limit both the content of the answer and the form of the written response. Multiple-choice interpretive exercises assess many abilities more reliably than restricted response essays. The advantages of restricted response are that



44  
this format narrows the focus of your assessment to a specific and well defined performance. The nature of these items makes it more likely that your students will interpret each question the way you intended. 21  
You are in a better position to assess the correctness of students' answers when a question is focused and all students interpret it in the same way. When you are clear about what makes up correct answers, your scoring ability, and hence the scorer's validity improves.

## 2.6 Requirements of a Good Test

9  
Bachman and Palmer (1996: 17-37) states that test usefulness provides a kind of metric by which we can evaluate not only the tests that we develop and use, but also all aspects of test development and use. There are six test qualities-reliability, construct validity, authenticity, interactiveness, impact and practicality. Another similar principle is stated by Browns (2004: 19-39): a good test has three classic criteria: practicality, reliability and validity. Practicality means it is within the means of financial limitations, time constraints, ease of administration, and scoring interpretation. The value and quality of a test are dependent upon such nitty-gritty, practical considerations. 20  
Practicality is also concerned with whether the test is designed to be norm-referenced or criterion referenced. In norm-reference tests, each test-taker's score is interpreted in relation to a mean-median, standard deviation, and/or percentile rank. The purpose in such tests is to place test-takers along a mathematical continuum in rank order. Typical norm-reference tests are standardized tests administered to large audiences, with results quickly disseminated to test takers. 25

80  
Reliability is concerned with consistency and dependability. If the same test is given to the same subjects on two different occasions, the test itself should yield similar results, it should have reliability. 19  
Scorer reliability is the consistency of scoring by two or more scorers. If very subjective techniques are employed in the scoring of a test, one would not expect to find high scorer reliability. 25  
If scoring directions are clear and specific as to the exact details the judge should attend to, then such scoring can be reasonably consistent and dependable. 38

38  
There are some factors that may contribute to the unreliability of a test: student-related reliability, rater reliability, test administration reliability and test reliability. Poor test administration also can result in unreliability of the test. The conditions such as street noise outside, photocopying variation, 5  
the amount of light in different parts of the room, variations in temperature, and even the condition of desks and chairs also affect reliability of a test. The last item for reliability is test reliability that is the nature of the test itself that is potential to

6 cause measurement of errors. If a test is too long, test-takers may become fatigued by the time they reach the latter items and respond correctly. As quoted by Brown, Mousavi (2002:21) mentions 79 the most common learner-related issue in reliability which is called 27 student-related reliability. This is caused by temporary illness, fatigue, a bad day, and other physical or psychological factor.

Another important factor that influences the quality of the test is validity. Browns (2004:2) states that validity covers 52 content-related evidence, criterion-related evidence, 4 construct-related evidence, consequential validity, and face validity. Validity is the extent to which a test calls for performance that matches the course or unit study being tested. It is also concerned with how well a test determines whether or not students have reached an established set of goals or level of competence.

17 Content related evidence refers to the samples of subject matter about which conclusions are to be drawn, and if it requires test-taker to perform the behavior that is being measured, it can claim content-related evidence. 6 Criterion related evidence is best demonstrated through a comparison of results of an assessment with results of some other measure of the same criterion.

5 Construct-related evidence is any theory, hypothesis, or model that attempts to explain observed phenomena in our universe or perceptions. Constructs may or may not be directly or empirically measured-their verification often requires inferential data.

Consequential validity 7 encompasses all the consequences of a test, including consideration of the accuracy in measuring intended criteria, its impact on the preparation of test-takers, its effect on the learner, and the intended and unintended social consequences of test interpretation and use.

Face validity refers to the degree to which a test looks right, and appears to measure the knowledge or ability it claims to measure, based on subjective judgement of the examinees who take it, the administrative personnel 28 who decide on its use, and other psychometrically unsophisticated observers. To get high face validity 6 learners encounter a well-constructed, expected format with familiar tasks, a test that is clearly achievable within the allotted time limit, items that are clear and uncomplicated, directions that are crystal clear, tasks that relate to their course work, and a difficulty level that presents a reasonable challenge (Mousavi, 2002 in Brown 2004:26).

4 The next principle of language test is authenticity. It refers to the degree of correspondence of the characteristics of a given language test task to the features of a target language task. In a test, authenticity may be present in: the language in the test is as natural as

possible, items are contextualized rather than isolated, topics are meaningful for the learner, some thematic organization to items is provided, such as through a storyline or episode, and tasks represent or closely approximate real-world tasks.

The last is <sup>27</sup>washback. It refers to the effects the tests have on instruction in terms of how students prepare for the test as well as how teachers teach. Washback effect <sup>5</sup>that occurs in the classroom assessment is the information that "washes back" to the students in the form of useful diagnoses of strengths and weaknesses. Washback also includes effects of an assessment on teaching learning prior to the assessment itself, that is, on preparation for the assessment.

## CHAPTER III PROCEDURES FOR TEST DEVELOPMENT

### 3.1 Purpose of Testing Reading Based on the Course Description

The test to be developed in the present report is a Reading test of an English course called Professional English that is offered to students in particular of Hotel Management Department in Diploma III in Tourism, Universitas Merdeka Malang. This means that logically the purpose of testing reading in the present report should be associated with achievement orientation of testing (Davies, 1990:20), in that the purpose of testing reading is aimed at knowing the learning progress or how far the students have learned the instructional objectives stated in course description.

As has been stated in the course outline, the purpose of giving reading skills in the course is formulated as follows (Course Outline: Professional English, ENG304): 'by the end of the course the students are expected to be able to read ... in hotel English with the ultimate goal in the sustainability in the professionalism in the hotel industry. In the present report, therefore, these will be the purpose of testing reading.'

### 3.2 Subjects of the Course

The subjects taking the course are those students with a total of 142 who are enrolled in Diploma 3 in Tourism, Merdeka University Malang in semester II in the academic year 2005-2006. There are 3 (three) parallel classes in all. These students got other English courses in the previous semesters. These are *Professional English* and *Communicative English* in semesters II and IV respectively. These two courses aim at providing the students with a focus on reading and writing skills in the case of *Professional English*; and a focus on speaking skills in the case of *Communicative English*. The kind of English taught in these two courses is directed to the mastery of functional English that is associated with the *rank and file* level or operator level. In addition to this, these students also have performed their on-the-job training in hotels where English is also used. In short, it can be assumed that the students have mastered basic English for daily communication functionally. This follows that they are familiar with basic English, which will facilitate their reading skills when they are posed with English specifically related with managerial matters.

### 3.3 Test Content Specification

The ultimate objective of the reading course - Professional English - to provide the students with the ability to read English texts of the intermediate to upper intermediate levels, related to tourism industries, particularly hotel-management matters. As such, the measurement of this course objective is addressed to the basic competence to be revealed through the reading test that is developed.

As has been discussed in Chapter II, theoretically, several important factors affect reading comprehension (Bachman, 1990; Grabe, 1991). Grabe (1991) outlines six factors: automatic recognition skills, vocabulary and structural knowledge, formal discourse structure knowledge, content/world background knowledge, synthesis and evaluation skills/strategies and meta cognitive knowledge and skills monitoring. In a more operational term, the function of these factors plays an essential role that enables reading comprehension to happen. Insufficiency or incompetence in one or more of these factors can hinder comprehension. In other words, if comprehension is understood as understanding text contents which requires the activation of those factors, then, the dimension of reading comprehension in this context is understanding text contents, particularly these texts of the intermediate to upper intermediate levels of a particular genre (Celce-Murcia, Dornyei, and Thurrel, 1995); that is, the ones related to tourism industries, particularly hotel-management matters.

Of the dimension, the functional competencies that a reader needs to have, which in Munby's terms are called micro reading skills (1978), are several. Other conceptualizations of functional reading competencies are also used in standardized proficiency tests like the TOEFL (Gear and Robert, 1996) as well as the IELTS (Alderson, 2000:131). The reading course in the present report is designed to include some of the functional competencies proposed by Munby (1978), the TOEFL (Gear and Robert, 1996) and the IELTS (Alderson, 2000:131) These include the following:

1. locating factual information
2. determining propositional inferences of a text
3. evaluating expressions in a text
4. summarizing and inferring underlying information in a text
5. understanding information in a text through visual presentations

Thus, these "micro skills" outlined above are the functional reading skills to be measured in the present report.



These functional reading skills are essentially conceptual in that these cannot be directly measured. For the purpose of measurement, these skills need to be derived in terms of indicators that reflect the essence of these functional reading skills. A critical analysis on each of these skills based on the ideas proposed by Wiener and Bazerman (1988) and Anderson, Durston and Poole (1969) yields indicators for each of these skills. The resulting indicators as determined in the reading course in the present report are as follows:

1. locating factual information
  - finding main factual information: *who, what, where, and/or when*;
  - finding factual explanatory: *why and how*
2. determining propositional inferences of a text
  - finding propositional informative inferences
  - finding propositional exploratory inferences
3. evaluating expressions in a text
  - identifying values of expressions in the forms of facts, opinions evidence, definition, or implication
4. summarizing and inferring underlying information in a text
  - drawing a logical conclusion,
  - predicting outcomes, and/or
  - generalizing
5. understanding information in a text through visual presentations
  - presenting information contained in a text through visual presentations like a picture, a chart, or a curve

A critical analysis attempted <sup>69</sup>for the purpose of the test content specification required for the test in the present report thus far has shown that the analysis has advanced from the very concept of reading ability as is envisioned in the course through the analysis of its dimension and micro skills of reading to very operational terms, that is the indicators of these micro skills. The outline of the test content specification for the purpose of testing the students' achievement in the course is outlined in Table 3.1 as follows:

Table 3.1: Test Content Specification

Basic Competence (BC)	Dimension (D)	Functional Competence (FC)	Performance Indicators (PI)
Ability to read English texts, of the intermediate to upper intermediate levels, related to tourism industries, particularly hotel-management matters	Understanding text content (UTC)	1. location factual information	<ul style="list-style-type: none"> <li>finding main factual information: <i>who, what, where</i> and/or <i>when</i>;</li> <li>finding factual eksplanatory : <i>why</i> and <i>how</i></li> </ul>
		2. determining propositional inferences of a text	<ul style="list-style-type: none"> <li>finding propositional informative inferences</li> <li>finding propositional explanatory inferences</li> </ul>
		3. evaluating expressions in a text	<ul style="list-style-type: none"> <li>identifying values of expressions in the forms of fact, opinions, evidence, definition, or implication</li> </ul>
		4. summarizing and inferring underlying information in a text	<ul style="list-style-type: none"> <li>drawing a logical conclusion,</li> <li>predicting outcomes, and/or</li> <li>generalizing</li> </ul>
		5. understanding information in a text through visual presentation	<ul style="list-style-type: none"> <li>presenting information contained in a text through visual presentations like a picture, a chart, or a curve</li> </ul>

In testing the students' achievement in the course, measurement of reading ability is actually addressed to deal with these performance indicators which logically reflect the achievement of their functional competencies as well as dimension and basic competence.

### 3.4 Selection of the Test Format

The next step to perform when the test content specification has been established is to select the format of the test to be used. Nitko (1996:124) cautions that three principles need exercising when crafting assessment tasks. These principles are (1) focusing merely on

essential learning objectives,(2) eliciting from the students abilities relevant to the learning objectives, and (3) neither preventing nor inhibiting students' ability to demonstrate achievement of learning objectives. In the selection of format for the test used in the course, these principles are as far as possible considered. For instance, to maintain principles 1 and 2, accurate analyses of test content have been established in the test content specification as shown in Table 3.1. Principle 3 seems most relevant for the purpose of selection of test format to be used.

Several testing formats may be used to test different reading abilities (Brown, 2004:186-216; Hughes, 2003:136-159; Douglas, 2000:189-245, Alderson, 2000:202-270; Djiwandono, 1996:62-67). The test format employed to measure students' reading achievement in the course is a short answer type (Brown, 2004:206; Nitko, 1996:124-128). According to Odell and Wesman, basically there are three types of short-answer format: question, completion and association. The question task asks direct questions; the completion one requires the test takers to add words to complete an incomplete statement; and association variety has a table or a diagram for which the test takers to recall numbers (Nitko, 1996:124). The first and the last formats are used in testing reading abilities.

There are several reasons for using the short answer task. This format is its flexibility in that it can be used to assess not only lower-order thinking skills but also higher-order levels of ability such as making simple interpretations of data and applications of rules, solving numerical problems, and manipulating and balancing mathematical symbols (Nitko, 1996:125). In addition to these, the format offers several practical and technical advantages as follows. The short answer task is practical in the construction as well as in the objective scoring. It offers a technical merit in that it can minimize the test takers' blind guessing. Also, while the format is expected to reveal the students' understanding of the text accurately, the format can minimize problems of expressing answers using their own words, which tends to '...add a significantly difficult task on top of reading (Hughes, 2003:143).

### **3.5 Text Selection**

Selecting a text is another critical stage that is carried out in the development of the reading test. It is said to be critical because the appropriate conduct of text selection will have the effect on level of readability of the text which further affects level of difficulty of the text. Ultimately, the validity of the test will also be affected (Cohen et. al, 1988:169-173). For the purpose, several points need to be established prior to text selection.



First, the test to be developed is an achievement test. As such, there is a desirable need to link between what is taught and how it is tested (Davies, 1990:20), including the use of the text in the best. As has been discussed in Parts I and II, the test is aimed at assessing reading abilities in English of students in particular of Hotel Management Department in Diploma 3 in Tourism, Merdeka University Malang. In addition, in the course these students are directed at the managerial level. Therefore, the text to be used should also be about hotel management matters at the managerial level. This strategy, it is argued, is to maintain the schemata and background <sup>109</sup> knowledge on the part of the test takers' variables that affect reading (Alderson, 2000:33-34) as well as the principle of testing of language for specific purposes (LSP) <sup>3</sup> (Douglas, 2000:1-2) stating that 'test content and test methods' are derived from a specific language use context.

As a consequence of the first consideration described in the section above, the texts to be selected are those derivable from sources that reflect the need. For the purpose, four sources are referred to: *The Japan Journal* (2005), *Tourism Today* (2002), *Introduction to Hospitality* by John R. Walker, and (1999) *Managing Hotels Effectively* by Eddystone C. Nebel III (1991). Besides being relevant to the needs of the test, these sources are also authoritative and actual in that they become important references to students and lecturers alike for matters related with tourism industries in general and hotel management in particular.

Procedures taken for the text selection is as follows. With a specific view to the basic competence (BC), dimension (D), functional competence (FC), and performance indicators (PI) as outlined in the test content specification, reading texts that reflect these aspects are to be selected. In addition, text parameters are also used as a basis for selecting texts. These <sup>43</sup> parameters are text types, text forms, graphic features, topics, styles, intended readership, length, readability, range of vocabulary, range of grammar (Hughes, 2003:140). The following table outlines the text parameters with their corresponding approximate criterial requirements.

Table 3.2: Approximate Requirements of Texts

No.	Text Parameters	Approximate Requirements
1	Text types	Textbooks, journal articles
2	Text forms	Descriptive, expository
3	Graphic features	Tables, charts, diagrams
4	Topics	Hotel management, tourism-related
5	Styles	Formal
6	Intended readership	Hotel-management students
7	Length	About 150 - 500 words
8	Readability	Like authentic texts in textbooks and journal
9	Range of vocabulary	Technical
10	Range of grammar	Compound, complex

Based on this process, 15 (fifteen) reading texts from those four sources are then identified. These texts are 13 (thirteen) in the form of short paragraphs and 2 (two) in the form of short essay. These 15 (fifteen) identified reading texts from the different sources are then taken out from their original sources and are rewritten as a pool of texts (See Appendix 1 for all the initial texts identified).

These 15 (fifteen) identified reading texts are further reviewed more closely and critically with respect to performance indicators (PI) of each functional competence (FC). The review is focused primarily on the suitability of the identified texts with indicators (PI) for the purpose of developing corresponding test items (Hughes, 2003:142). For example, in order to satisfy the performance indicator 'finding main factual information: *who*, *what*, *where*, and/or *when*', the text should have contents that enable the potential items to assess factual information: *who*, *what*, *where*, and/or *when*. Based on the review, out of the 15 (fifteen) identified reading texts, 8 (eight) reading texts are selected as the reading texts to be used in the test. These texts then serve as the input texts for the test items to be established.

### 3.6 Item Writing and Item Assembling

Item writing includes two important parts: the writing of directions and that of the test items themselves. The following is the description of each. In writing test directions, the first thing to do is to examine the common characteristics that all the items in all the 8 (eight) reading texts will share. As there are 8 (eight) reading texts, the information concerning the number of the text needs also informing to students. Also, as short answers are required, information on how to accomplish the items is also provided. Finally, to ensure that the

students can demonstrate their reading abilities, caution not to cooperate during the testing session and not to use dictionary are also included.

During the writing of the test items, three general aspects are exercised. The items are necessarily those that (1) measure the abilities as outlined in the test content specification. (2) draw out reliable reading behaviors from the students, and (3) allow for reliable scoring to be performed (Hughes, 2003:143).

The writing of the test items is based on a procedure that follows. To really understand the text content, it is first necessary to have a very careful reading of the texts selected. This reading will provide a clear picture of the scope of contents of the reading texts that are potential for questions to be raised. In this stage, notes to highlight aspects to be tested are put below each reading text. These include notes on main ideas, relevant details, signals of transitions, stages of text development etc. The next step to be taken is to consider the type of the task that is realistic to expect the <sup>78</sup>test takers to be able to act upon the test items. When the test task is decided, then, it is time to write the draft of the test items. The final stage to be carried out is to review the accuracy of the language used in the items and the suitability of the draft of the test items already written with regards to the intended performance indicators.

When all the intended indicators are already measured using their corresponding test items, and the items reviewed for language accuracy and validity, the items are then assembled into a set of test. The strategy used to assemble the test items is based on the assumed level difficulty of the items. Thus, the items are arranged from the easy to the difficult ones, following their corresponding reading texts. Besides, test identity that includes date, time of administration, etc, and directions <sup>14</sup>for doing the test are also considered. The plan of the item placement in the test is presented in Table 3.3 as follows.

Table 3.3: Final Item Placement<sup>1</sup> and Their Abilities Referred to

Basic Competence (BC)	Dimension (D)	Functional Competence (FC)	Performance Indicators (PI)	Item
Ability to read English texts, of the intermediate to upper intermediate levels, related to tourism industries, particularly hotel-management matters	Understanding text contents (UTC:	1. locating factual information	<ul style="list-style-type: none"><li>finding main factual information <i>who, what, where, and/or when</i>;</li><li>finding factual explanatory: <i>why</i> and <i>how</i></li></ul>	1, 3, 7, 8, 9, 10, 11, 23, 27
		2. determining propositional inferences of the text	<ul style="list-style-type: none"><li>finding propositional informative inferences</li><li>finding propositional exploratory inferences</li></ul>	4, 5, 16, 22, 31
		3. evaluating expressions in a text	<ul style="list-style-type: none"><li>identifying values of expressions facts, opinions, evidence, definition, or implication</li></ul>	17, 18, 15, 19, 20
		4. summarizing and inferring underlying information in a text	<ul style="list-style-type: none"><li>drawing a logical conclusion,</li><li>predicting outcomes, and/or</li><li>generalizing</li></ul>	2, 21, 12, 13, 14, 24, 25, 26, 28, 30
		5. understanding information in a text through visual presentations	<ul style="list-style-type: none"><li>presenting information contained in a text through visual presentations like a picture, a chart, or a curve</li></ul>	6, 29
Total of the items				31

<sup>1</sup> Applicable only after revisions based on moderators' feedback and informal trail

### **3.7 Writing the Key Answers and Test Scoring Guide**

When the reading test that consists of texts with their corresponding items is completed, two further important aspects of the test need to be established. These are the test key answers and test scoring guide. The following is the description of each. In providing test key answers, the principle of short answer type is maintained. In the direction, the students are to provide an answer as short as possible to the question not more than two sentences at the maximum. Therefore, the key answer is made on this basis.

The provision of test key answers is based on the procedures that follow. First, each reading text is carefully read once or twice to ensure that the content is understood correctly. Then, each question following the text is tried with an answer as a result. The expected answer to each item is then recorded as a basis of key answers of the test. It is expected that with the key answers at hand, subjectivity of the answers of the test is minimized, thus maintaining reliability of the test (Brown, 1996:185-223). Based on the principle described above, the key answer to each item is provided (see Appendix 9).

The scoring system employed is based on the dichotomous scoring principle. A correct answer based on the key answer is assigned a score of 1 (one) and an incorrect one has a score of 0 (zero). However, weighting is employed, particularly with regards to the value of each of the functional competencies referred to. It is desirable to recall that there are 5 (five) functional competencies as outlined in the test content specification. These are (1) locating factual information, (2) determining propositional inferences of a text, (3) evaluating expressions in a text, (4) summarizing and inferring underlying information in a text, and (5) understanding information in a text through visual presentations. Assuming that there is an increase in the level of difficulty from functional competence 1 to functional competence 5, then, the weighting to the score obtained in each of these five functional competencies is assigned intuitively, in the sense that the weighting increases by one point from functional competence 1 'locating factual information' to functional competence 5 'understanding information in a text through visual presentations'. Thus, a correct answer obtained in all items in functional competence 1 'locating factual information' is weighted by 1 (one) point, a correct answer obtained in all items in functional competence 2 'determining propositional inferences of a text' by 2 (two) points, a correct answer obtained in all items in functional competence 3 'evaluating expressions in a text' by 3 (three) points, a correct answer obtained in all items in functional competence 4 'summarizing and inferring underlying information in a text' by 4 (four) points, and a correct answer obtained in all items in functional competence

5 'understanding information in a text through visual presentations' is weighted by 5 (five) points. The plan of the scoring scheme with weighting is presented in Table 3.4.

As an illustration, if item 1 is correctly answered, the score obtained will be 1 (1 times 1 = 1). A correct answer of item 42 will yield a score of 2 (1 times 2 = 2). A test taker doing item 40 which is answered correctly will have a score of 5 (1 times 5 = 5), and so on.

Table 3.4: The Scoring Scheme

Functional Competence (FC)	Performance Indicators (PI)	Item	Weight Value
1. locating factual information	<ul style="list-style-type: none"> <li>finding main factual information: <i>who</i>, <i>what</i>, <i>where</i>, and or <i>when</i>;</li> <li>finding factual explanatory: <i>why</i> and <i>how</i></li> </ul>	1, 3, 7, 8, 9, 10, 11, 23, 27	1
2. determining propositional inferences of a text	<ul style="list-style-type: none"> <li>finding propositional informative inferences</li> <li>finding propositional exploratory inferences</li> </ul>	4, 5, 16, 22, 31	2
3. evaluating expressions in a text	<ul style="list-style-type: none"> <li>identifying values of expressions facts, opinions, evidence, definition or implication</li> </ul>	17, 18, 15, 19, 20	3
4. summarizing and inferring underlying information in the text	<ul style="list-style-type: none"> <li>drawing a logical conclusion,</li> <li>predicting outcomes, and/or</li> <li>generalizing</li> </ul>	2, 21, 12, 13, 14, 24, 25, 26, 28, 30	4
5. understanding information in a text through visual presentations	<ul style="list-style-type: none"> <li>presenting information contained in a text through visual presentations like a picture, a chart, or a curve</li> </ul>	6, 29	5

### 3.8 Grading and Grade Interpretation

Dirwandono (1996:117) outlines that an appropriate evaluation towards the students' performance in a test necessitates 2 (two) critical phases. The first phase is examination into the students' work to determine whether the students' response is correct or incorrect according to the key answer and the scoring scheme used. The output of this process is raw scores. This process may involve the use of numerical figures. All the procedures that are described in section 3.7 Writing the Key Answers and Test Scoring Guide above include this phase. The second important phase needed to be considered is grading. This is essentially a process of indicating the students' standing in relation with a particular reference in which grade interpretation is involved. The following is the description of grading and grade interpretation.

#### 3.8.1 Grading

Based on the scoring scheme outlined in section 3.7 above, it can be stated that the score that a test taker may obtain is the function of the correct answers s/he can get and the weighting of the correct answers. The table that follows may be used to clarify this scoring.

Table 3.5: Score Weighting Calculation<sup>1</sup>

Item	Weight Value	Grade
1, 3, 7, 8, 9, 10, 11, 23, 27	1	9 x 1 = 9
4, 5, 16, 22, 31	2	5 x 2 = 10
17, 18, 15, 19, 21	3	5 x 3 = 15
2, 21, 12, 13, 14, 24, 25, 26, 28, 30	4	10 x 4 = 40
6, 29	5	2 x 5 = 10
Maximum Total Grade		84

As shown in the table above, a student getting a correct answer each with items 1, 3, 7, 8, 9, 10, 11, 23, and 27 will get a score 1 with a total score 9 points; a student getting a correct answer each with items 4, 5, 16, 22, and 31 will get a score 2 with a total score 10 points; and so on. With this procedure of scoring, a maximum total grade that a student may obtain is 84 (eighty four) points and a minimum grade 0 (zero). That maximum score, 84

<sup>1</sup> Applicable after revisions based on moderators' feedback and informal trial



(eighty four) points, indicates the highest level of mastery of the students in reading skills up to the mid-term period which is a half of the course progress as designated in the course description. In the same way, that minimum score, 0 (zero) point, indicates the absence of mastery of the students in reading skills.

### 3.8.2 Grade Interpretation

Interpretation of students' grade is one important step in testing and evaluation (Djiwandono, 1996:120-127; Nitko, 1996:349-406). Interpreting students' score can be based primarily on the principle of the norm-referenced approach or the criterion-referenced approach (Lynn and Gronlund, undated:439-466; Bachman, 1990:210-214; cf. Brown, 1996:2-8).

As there are approaches to interpretation of grades, it is essential that the approaches be appropriately understood so that the interpretation of the scores to be performed can be conceptually based and accurately carried out. Lynn and Gronlund, (undated:443) state that <sup>41</sup> '...norm-referenced test interpretation tells us how an individual compares with other persons who have taken the same test' and <sup>7</sup> '...criterion-referenced test interpretation permits us to describe an individual's test performance without referring to the performance of others...' who have taken the same test but it enables the description of an individual's test performance against 'well-defined and fairly specific objectives' (Brown, 1996:2). The definition clearly shows the different underlying principle that the approaches hold. Brown (1996:3) further outlines clearly several differences of the tests based on those two approaches, three of which are fundamental. First, in terms of type of interpretation, tests that are norm-referenced are relative in that a student's performance in the test is compared to that of all other students in percentile terms. On the other hand, tests that are criterion-referenced are absolute in that a student's performance in the test <sup>65</sup> is compared only to the amount, or percentage, of material <sup>35</sup> learned. Next, norm-referenced tests are designed to measure general language abilities or proficiencies whereas criterion-referenced test are to specific objectives-based language points. Finally, the purpose of using norm-referenced tests is to spread the students out along a continuum of general abilities or proficiencies; meanwhile, criterion-referenced tests aim at assessing the amount of instructional materials learned by individual students. <sup>108</sup>

The present test is designed to assess the students' learning progress in reading skill in the course, particularly their achievement up to the mid-term period. As such, the approach that is appropriate to interpret the students' scores or grades is that of criterion-referenced.

The criteria referred to in the test are then the functional competencies as reflected in their performance indicators (see Table 3.1 Test Content Specification). The students' performance in the test which is shown by their scores or grades reflects the level of the amount of mastery towards the overall learning objectives stated in the functional competencies, expressed in percentage.

In order to conform the points of evaluating the students' achievement as ruled in *Pedoman Penyelenggaraan Pendidikan Tahun Akademik 2004/2005, Universitas Merdeka Malang* in page 18, those maximum and minimum levels of test performance (being 84 and 0 points) need to be transformed linearly. This is performed by employing the following formula:  $CA:84 \times 100\%$ , where CA = obtained grade . Referring to Criterion Reference Evaluation as *Pedoman Penyelenggaraan Pendidikan Tahun Akademik 2004/2005, Universitas Merdeka Malang* aspires, this formula still clearly maintains the students' level of mastery as is shown by the use of percentage. With this procedure in mind, the students' obtained grade then can be assessed against performance criteria set up in *Pedoman Penyelenggaraan Pendidikan Tahun Akademik 2004 2005, Universitas Merdeka Malang* as follows:

Table 3.6: Grade Conversion Table

Transformed Score	Grade
80 - 100	A
66 - 79	B
56 - 65	C
45 - 55	D
0 - 44	E

Adapted from *Pedoman Penyelenggaraan Pendidikan Tahun Akademik 2004/2005, Universitas Merdeka Malang*, p. 18.

### 3.9 Test Item Moderation

With a view that item writing that has been completed yields a first rough draft of the intended test with its test items considered provisional, moderation applied to test items constitutes another important step in test development (Djiwandono, 1996:112). Moderation is '...the scrutiny of proposed items' written by the test developer that is performed by (ideally) at least two colleagues, neither of whom is the author of the items being examined'

(Hughes, 2003:63). This step is essentially 'further validation check' (Weir, 1990:39-40), the purpose of which is critically to find weaknesses in the test items already written. These weaknesses may include suitability of texts, format, and items (Weir, 1990:39-41).

### 3.9.1 Moderation Processes

With the concept described above, moderation in the present test is performed using the following procedures. The first stage taken is self moderation. This procedure is the one in which the present developer of the test under examination had a closer and critical look at the items that she had already developed. Djiwandono (1996:112) states that this procedure is necessarily avoided and just constitutes a minimal attempt made to moderate test items. This is true if there is no further quality check on the items, that is, the test is administered for use immediately upon item writing. In the present test development, however, this self-moderation represent the first stage of other stages taken for test item moderation.

Self moderation was concentrated mainly at two aspects. These are lexical and grammatical accuracy in the writing of the sentences and suitability of the items to measure their corresponding performance indicators. The former was carried out by rereading all the items very carefully; the latter by checking the sentential formulation of the test items with regards to their corresponding performance indicators. It turned out that some necessary revisions were made on some items in this stage.

The next stage is moderation by colleagues. In this stage, a panel consisting of 3 (three) test item moderators was involved. All of them were lecturers of the course - Professional English - who have extensive English teaching experiences ranging from 5 to 25 years. Two of the moderators hold a *sarjana* (S-1: Dra) degree in English education and one has a *magister's* (M.Pd) degree. It is believed that these three moderators had a sufficient capacity to play their role critically as the present-test moderators.

The examination on the test items was emphasized on 9 (nine) aspects as follows: agreement between the performance indicator and the item to measure it, the wording of the item, function of the measurement, order of presentation, instruction, level of difficulty, time allotment, and layout. The moderators were to assess all the items on these aspects against categorical judgment in a three-point scale. In addition to these, the moderators were also to put critical notes directly on items they thought needing commenting.

For the purpose of moderating test items, the moderators were provided with a review sheet containing 9 (nine) aspects to assess the test items: agreement between the performance

indicator and the item to measure it, the wording of the item, function of the measurement, order of presentation, instruction, level of difficulty, time allotment, and layout with their three-level categorical scale of measurement (see Appendix 3 for the sheet), test content specification, and the reading test with its items (see Appendix 2 for the test draft for moderation purposes). The moderators were to perform the moderation of the items for about 2 (two) weeks in October 2005.

On the 9 (nine) aspects to be assessed of the test: agreement between the performance indicator and the item to measure it, the wording of the item, function of the measurement, order of presentation, instruction, level of difficulty, time allotment, and layout, a set of criteria of acceptance is established. These criteria are:

1. most or all of the moderators should rate 'good' on the aspect 'agreement between the performance indicator and the item to measure it';
2. most or all of the moderators should rate 'clear' on the aspect 'wording of the item';
3. most or all of the moderators should rate 'good' on the aspect 'function of the measurement';
4. most or all of the moderators should rate 'logical' on the aspect 'order of presentation';
5. most or all of the moderators should rate 'clear' on the aspect 'instruction';
6. most or all of the moderators should rate 'fair' on the aspect 'difficulty';
7. most or all of the moderators should rate 'right' on the aspect 'time allotment';
8. most or all of the moderators should rate 'good' on the aspect 'layout'.

These criteria are used as a basis for making necessary relevant revisions of the test that has been made.

Along with the colleague moderation, moderation by an expert was also performed. An expert in language testing was involved in this stage. The moderator held a doctorate degree in English language education. He has ample and intensive experiences in English teaching and in language testing as well as English language test development. Unlike those three moderators, this moderator, however, was not provided with a review sheet containing 9 (nine) aspects to assess the test items. Rather, he was given the test content specification, and the reading test with its items to be commented on.

### 3.9.2 Moderators' Feedback and Its Analysis

There are 2 (two) important points to be considered in examining the moderators' responses. The first point is to ensure that the moderators are consistent in their responses. Secondly, the moderators' responses satisfy the criteria for revisions of the test described in the previous section. The following is the description of each.

The feedback provided by the three moderators' results in the summary that is presented in Table 3.7.

Table 3.7: Summary on Moderator's Feedback

Aspects	Moderator's Response		
	Moderator 1	Moderator 2	Moderator 3
agreement	2	3	3
wording	3	3	3
function	2	3	3
order	3	2	3
instruction	2	2	2
difficulty	2	2	2
time allotment	1	2	1
layout	3	2	1

To examine the consistency of responses among moderators, the table for moderators' response analysis is required as shown in Table 3.8.

Table 3.8: Table Analyzing Moderator's Feedback

Aspects	Moderator's Response			$X_i$	$X_i^2$
	Moderator 1	Moderator 2	Moderator 3		
agreement	2	3	3	8	64
wording	3	3	3	9	81
function	2	3	3	8	64
order	3	2	3	8	64
instruction	2	2	2	6	36
difficulty	2	2	2	6	36
time allotment	1	2	1	4	16
layout	3	2	3	8	64
$X_i$	18	19	20	57	425
$X_i^2$	324	361	400	1085	

Note:

$X_i$  and  $X_i$  = total scores of each case;

$X_i^2$  and  $X_i^2$  = total of each case's squared score

To examine the consistency among the moderators (r), the following formula is employed:

$$r = \frac{MSS_a - MSS_e}{MSS_a} \quad (3.1)$$

where:

r = the consistency among the moderators

MSS<sub>a</sub> = mean of sum of square of aspect

MSS<sub>e</sub> = mean of sum of square of error

To compute the value of  $MSS_a$  (mean of sum of square of aspect), the following formula is employed:

$$MSS_a = \frac{SS_a}{df_a} \quad (3.2)$$

where:

$SS_a$  = sum of square of aspect

$df_a$  = degree of freedom of aspect.

$SS_a$  (sum of square of aspect) is determined by

$$SS_a = \frac{1}{N_m} (\sum X_i^2) - \frac{(X_i)^2}{N_{ma}} \quad (3.3)$$

$SS_m$  (sum of square of moderator) is determined by

$$SS_m = \frac{1}{N_a} (\sum X_j^2) - \frac{(X_j)^2}{N_{ma}} \quad (3.4)$$

and  $SS_t$  (sum of square of total) is determined by

$$SS_t = (\sum x^2) - \frac{(X_i)^2}{N_{ma}} \quad (3.5)$$

where:

$x$  = cases

$N_{ma} = N_m \times N_a$

To compute the value of  $MSS_e$  (mean of sum of square of error), the following formula is employed :

$$MSS_e = \frac{SS_e}{dfe} \quad (3.6)$$

where:

$SS_e$  = sum of square of error =  $SS_t - SS_a - SS_m$  (3.7)

$dfe$  = degree of freedom of error =  $df_t - df_m - df_a$  (3.8)



Based on the computation using the formulas presented previously (3.1 through 3.8) (see Appendix 13 for the complete manual computation), the values of each parameter can be summarized as Table 3.9 that follows:

Table 3.9: Summary of Statistics for Moderators' Consistency

SS <sub>a</sub>	SS <sub>m</sub>	SS <sub>t</sub>	SS <sub>e</sub>	Df <sub>a</sub>	df <sub>m</sub>	df <sub>t</sub>	MSS <sub>a</sub>	MSS <sub>e</sub>
6.291	.25	9.625	3.083	7	2	23	.899	.237
				(8-1)	(3-1)	(24-1)		

Thus, the coefficient of consistency among the moderators (r) can be computed:

$$\begin{aligned}
 r &= \frac{MSS_a - MSS_e}{MSS_a} \\
 &= \frac{.899 - .237}{.899} \rightarrow = .635
 \end{aligned}$$

According to Brown (1996), the value of the coefficient of consistency which figures at .635 among the moderators (r) indicates that moderators involved were sufficiently consistent in scoring the aspects to be assessed, like agreement between the performance indicator and the item to measure it, the wording of the item, function of the measurement, order of presentation, instruction, level of difficulty, time allotment, and layout.

Further, to examine the degree of agreement among the moderators in responding to each of the aspects to be examined, frequency of occurrence of the moderators' responses on each aspect to be rated was recorded. Table 3.10 records the moderators' responses.

Table 3.10: Agreement of the Moderator's Responses

Aspects	Scale to Choose	Frequency	Percentage	Follow-up Notes
Agreement between Items and Performance Indicators	1. unrelated	0	0%	no revision
	2. fair	1	33.3%	
	3. close	2	66.7%	
Item Wording	1. not clear	0	0%	no revision
	2. fair	0	0%	
	3. clear	3	100%	
Test Function	1. bad	0	0%	no revision
	2. fair	1	33.3%	
	3. good	2	66.7%	
Order of Presentation	1. bad	0	0%	no revision
	2. fair	1	33.3%	
	3. good	2	66.7%	
Instruction	1. not clear	0	0%3	revision
	2. fair	3	100%	
	3. clear	0	0%	
Difficulty	1. easy	0	0%	no revision
	2. moderate	3	100%	
	3. difficult	0	0%	
Time Allotment	1. little	2	66.7%	revision
	2. right	1	33.3%	
	3. much	0	0%	
Test Layout	1. bad	0	0%	no revision
	2. fair	1	33.3%	
	3. good	2	66.7%	

The table shows that aspects 'agreement, wording, function, order, difficulty and layout' satisfy the criteria; whereas aspects instruction and time do not. Therefore, these findings become the focus of revising the aspects. Beside these findings, moderators' comments on the items are also sources of information that can be used as a basis of revisions. Therefore, their notes are also considered. For example, one moderator commented that some items indicated by her need rewording.

The expert moderator commented that the focus of the test needs to be placed solely on the measurement of students' understanding of the text content. Measurement of other aspects like students' understanding of text structure may be put aside in the test. This suggestion seems plausible. However, since the measurement of the indicators relates to the course objectives, omission of items that are considered measuring text structure will certainly

change the course objectives, and this is not expected at least for the present time during the test development because the teaching-learning process that had been conducted up to the time of the mid term examination had been oriented to achievement of the objectives. In order not to totally leave the objectives, therefore, some of the items considered measuring text structure aspects more heavily were omitted, others considered and reworded.

### 3.10 Test Item Revision

Based on the feedback provided by moderators, be it derived from the results of computational analyses or verbal comments, revisions were made. Based on computational analyses, attention was paid to aspects 'instruction' and 'time'. This follows then that strategies to revise these aspects were established. First, the test was reviewed on the spots based on the feedback. Then, parts that are considered weaknesses were identified. Next, revisions on the parts shown were attempted. However, where revisions seemed to be impossible to be made, the points were left out not to be included in the test.

Changes on the wording of instruction were made as shown in Table 3.11 as follows:

Table 3.11: Revisions on Instructions

Original Wording (Pre Moderation)	Revision
1. There are 9 short texts in this test. Do all the test Items following these texts;	1. There are 8* short texts in this test followed with 34 question in all.
2. Write your answer in the answer sheet provided;	2. Do all the 34 test items following these 8 texts by answering all the questions after each text, or by filling in the blank spaces provided in the task;
3. Give as brief answer as possible, max. two sentences;	3. Write your answer in the answer sheet provided;
4. Work independently and no dictionary is allowed to be used.	4. Answer as briefly as possible, max. one short sentence;
	5. Work by yourself;
	6. Do not open your dictionary.

Note: \*considering also moderator's feedback on time allotment

With regards to feedback on time allotment, it was decided that there was a need to ease the number of the items. In so doing, there was a problem concerning which items to be left out. A decision was then made to omit Text 3 together with its three corresponding questions. The reason is that as suggested by the expert moderator, this text is heavily designed to test text structure. As a result, of the 9 original texts in the test, there is 8 (eight) reading texts used as input.

Still in the concern of time allotment, suggestions made by the expert moderator not to include items that measure text structure were used as a basis to reduce the number of items. These items taken out of the test were items 2 and 4 of Text 1, items 6, and 7 of Text 2. Thus, there were 7 (seven) items that were left out from the test. After revisions, the number of existing items used in the test is 35 (thirty five).

Beside omission of the number of the items, revisions, which were based on the verbal feedback made by moderators, were also worked on wording of the items. The changes made are shown in Table 3.12.

Table 3.12: Revisions on Item Wording

Original Wording (Pre Moderation)		Revision
Source	Questions	
Text 1	<ul style="list-style-type: none"> <li>What does the writer actually say about hotels in the text above?</li> </ul>	<ul style="list-style-type: none"> <li>What is the writer's opinion about hotel in the text above?</li> </ul>
	<ul style="list-style-type: none"> <li>Why does the writer use his/her personal experience in attending a gala banquet?</li> </ul>	<ul style="list-style-type: none"> <li>What is the purpose of the writer to present his/her personal experience in attending a gala banquet?</li> </ul>
Text 2	<ul style="list-style-type: none"> <li>According to the writer, what responsibilities do managers need to have?</li> </ul>	<ul style="list-style-type: none"> <li>According to the writer, what are the responsibilities of managers?</li> </ul>
	<ul style="list-style-type: none"> <li>What is meant by <u>basic principles</u> in the text?</li> </ul>	<ul style="list-style-type: none"> <li>What does the writer mean by <u>basic principle</u> (line 3 from bottom) in the text?</li> </ul>
	<ul style="list-style-type: none"> <li>Provide an example of a case where a manager works together with staff members on the job!</li> </ul>	<ul style="list-style-type: none"> <li>A manager need to work together with staff members on the job. Give an example!</li> </ul>

Original Wording (Pre Moderation)		Revision
Source	Questions	
Text 5	<ul style="list-style-type: none"> <li>What the foreign influence is there existing in both <i>soba</i> and <i>udon</i>?</li> </ul>	<ul style="list-style-type: none"> <li>What foreign variants effect both <i>soba</i> and <i>udon</i>?</li> </ul>
	<ul style="list-style-type: none"> <li>unlike <i>shabu-shabu</i> or <i>teriyaki</i>, why do we, foreigners, not recognized <i>nabe</i> as Japanese food)</li> </ul>	<ul style="list-style-type: none"> <li>Why do we, foreigners, not recognize <i>nabe</i> as food like <i>shabu-shabu</i> or <i>teriyaki</i>?</li> </ul>
Text 6	<ul style="list-style-type: none"> <li>What profession do you think is the writer of the text?</li> </ul>	<ul style="list-style-type: none"> <li>What is the profession of the writer of the text?</li> </ul>
	<ul style="list-style-type: none"> <li>What for did the writer spend time with management of good hotels?</li> </ul>	<ul style="list-style-type: none"> <li>What is the purpose of the writer by spending time with management of good hotels?</li> </ul>
	<ul style="list-style-type: none"> <li>What do you think is lacking in the hotel with poorly managed hotels?</li> </ul>	<ul style="list-style-type: none"> <li>What is lacking in the hotels that are not managed well?</li> </ul>
Text 7	<ul style="list-style-type: none"> <li>Mention sentences in the text above that have words or phrases that reflect facts!</li> </ul>	<ul style="list-style-type: none"> <li>Write sentences or phrases in the text that contain facts?</li> </ul>
	<ul style="list-style-type: none"> <li>Which sentences in the text about that have a words or phrases that reflect opinions?</li> </ul>	<ul style="list-style-type: none"> <li>Which sentences or phrases in the text that reflect the writer's opinion?</li> </ul>
	<ul style="list-style-type: none"> <li>According to the text above, there are factors that make successfull leaders. Mention the factors according to general managers (GMs)!</li> </ul>	<ul style="list-style-type: none"> <li>According to general managers (GMs) in the text above, what factors make successful leaders?</li> </ul>
	<ul style="list-style-type: none"> <li>According to Frank Anderson, a department head who is a quiet person will be successful in his profession. Give your comment!</li> </ul>	<ul style="list-style-type: none"> <li>Do you agree with Frank Anderson that a department head who is a quiet person will be successful in his profession? Give your comment!</li> </ul>

### 3.11 Informal Tryout

Prior to formal tryout, the test that had been revised based on the feedback provided by moderators was then put into an informal administration as test pilot testing. The section that follows describes administration of the informal tryout, analysis of the results of the informal tryout and the follow-up measures after the informal tryout.

#### 3.11.1 Administration of the Informal Tryout

The subjects involved in the informal trial were 6 (six) <sup>31</sup> students of the English Department, Faculty of Culture and Humanities, State Islamic University of Malang. These <sup>77</sup> students were in their fifth semester. Thus, they were the third year students. These students were selected at random out of the 30 students in the class. By design in the department, those students had got all their English skills courses and English component courses. Therefore, by the time the informal trial was conducted, they <sup>31</sup> were assumed to have had sufficient mastery of the basics of English.

Compared to the real target audience of the test being developed, those sample students had more credits of English courses. Therefore, although they were not students of a tourism department, in terms of their English learning, they were assumed to have capabilities to understand technical texts of tourism and hotel matters and answer questions contained in the test. As such, those students were believed to resemble the target of the test being developed in terms of their English.

The informal trial took place in Room H.213, State Islamic University of Malang from 14.00 to 15.30 on 22 November 2005. To the sample students, the test tasks and answer sheets were distributed. During the informal test administration, the present test developer took notes significant occurrences pertinent to the informal trial attempt. When the time was over, a brief conference intended to dig up further information concerning the sample's opinion and feelings on the test was held for about 15 (fifteen) minutes. The focus was put on aspects like clarity of instruction, level of difficulty, time allotment, and other significant comments.

#### 3.11.2 Analysis of Informal Tryout Results

The results of the test informal trial were analyzed in terms of two categories: test-related information and conference-based data. The following is the description of each.

Of test-related information, two sources are recognized: one is the significant occurrences during the informal trial and the other deals with the students' responses which are contained in the answer sheet. During the test, it was observed several important points as follows. First, no one of the sample respondents asked unclear points in the test. It seemed that they could understand the instructions well. It could also mean that the questions or the test tasks were understandable. Next, apart from the attempt, all of the respondents did not manage to use the time allotted to try all of the items. This means that the time allotted was not sufficient. This is further evidenced from their response in the answer sheet that, of the six students, only two tried all the numbers in the test.

Further analytical examinations on their responses in the answer sheet revealed several findings as follows. In the first place, despite the instruction to write a short answer to each test task, all the samples still attempted most items in longer expressions which are not expected of them to do so. Next, out of 35 items, only one of the sample respondents scored 15 (fifteen) points as the correct answers (44.11%); two scored 11 correct answers (32.35%); one made 9 (nine) correct answers (26.47%); one answered 7 (seven) number correctly (20.59%); and one only made 2 (two) correct answers (5.88%). Then, of the items attempted by the respondents, 4 (four) items are answered correctly only by 5 (five) respondents. Another 4 (four) items are answered correctly only by 4 (four) respondents. Even, 10 (ten) items are answered wrongly by all respondents, meaning that no one could answer these ten items correctly (see Appendix 6 for the complete analysis of their response).

During a brief conference, two modes of data collection were performed. First, each of the sample respondents was asked questions concerning several aspects of the test: instruction, the wording of the test items, level of difficulty of items, time allotment, layout, presentation or gradation, and effect. To them, a brief questionnaire was distributed to be responded to according to what they felt during the administration (see Appendix 5). Results of the data collected after the administration of the trial are summarized in Table 3.13.

As shown in Table 3, 13, it was clear that to most (83.3%) the test instruction and test layout were fairly clear and fair; only one (16.7%) considered it not clear and bad respectively. As with test item wording and the grading of items based on difficulty level, most (83.3%) also considered them fairly clear and fair respectively. In terms of test item difficulty, all (100%) considered it moderate. To most (66.67%), the test items were challenging. It seems that time allotment is not agreeable to most (66.67%). In general, it can be inferred that excepting time allotment, the test posed no serious problems to the respondents.



In the second part during the conference, the sample respondents were given freedom to express what they felt about the test. They expressed their opinions that they felt the language (grammar and vocabulary) was not difficult for them to understand; however, they did not have clear ideas about the topics in most of the texts in the test as the texts, they claimed, were of a specific genre i.e. of hotel and tourism contexts. In addition, it turned out that students faced a problem in drawing a diagram and a curve particularly for responses required of them in items 7 and 33 respectively.

Table 3.13: Post-Informal Trial: Students' Feedback

Aspects	Scale to Choose	Frequency	Percentage
Instruction	1. not clear	1	16.7%
	2. fair	5	83.3%
	3. very clear	0	-
Wording	1. not clear	0	-
	2. fair	5	83.3%
	3. clear	1	16.7%
Difficulty	1. easy	0	-
	2. moderate	6	100%
	3. difficult	0	-
Time Allotment	1. short	4	66.67%
	2. right	2	33.33%
	3. long	0	'
Layout	1. bad	1	16.7%
	2. fair	5	83.3%
	3. good	0	-
Gradation	1. bad	0	-
	2. fair	5	83.3%
	3. good	1	16.7%
Effect	1. distressing	1	16.7%
	2. so and so	1	16.7%
	3. challenging	4	66.67%

Based on the analysis, it can be concluded that, first, the respondents tended to giving a rather long answers, which caused them to have language problems in answering the items. Also, some items pose non-technical problems to be accomplished; that is, a problem in drawing a diagram and a curve. Next, in terms of contents in general the test was difficult to the respondents. This explained another factor why the time allocated to was not enough. The factor that may cause the situation to be problematic is probably that the respondents were not familiar with the topics of the texts, which to them were mostly rather technically specialized.

### 3.11.3 Follow-up Steps

As shown by the results of the trial, the test was considered to be difficult to most of the respondents of the informal trial in terms of content as well as practices of completing some test tasks. Nevertheless, in terms of contents the test was believed to be close to the genre of hotel and tourism contexts, thus not posing a serious content-schemata problem to the real target of audience; that is, the students of hotel management department. This belief was based on the confessions made by the respondents during a brief conference after the administration of the informal trial that most were unfamiliar with the topics in the text.

Considering these, follow-up steps needing to be taken were decided as follows: (1) keeping the time allotment but reducing the items considered to be difficult based on the trial (particularly those items answered incorrectly by all respondents) with regards essentially to performance indicators outlined in the test content specification, however; (2) including in the answer sheets a flowchart and a curve scheme so as to avoid the students from technicalities of drawing or making such a flowchart or a curve during the test; (3) giving an example of how to mean by 'a short answer' in the answer sheet. The results of these strategies were a set of the final test comprising 31 items (see Appendix 7) and a revised form of answer sheet (see Appendix 8). These points are considered important to be attended to by the present test developer for the improvement of the test in general for the improvement of the test.

## **CHAPTER IV**

### **READING TEST TRYOUT AND RESULTS**

#### **4.1 Reading Test Administration**

Reading test administration basically constitutes the real field testing of the test that is developed to assess students' reading skills in the course Professional English IV. This section therefore reports the process of the administration of the reading test. Two main topics are thus presented: place and time of administration and description of test takers. The following is the description of each topic.

##### **4.1.1 Place and Time of Administration**

The test administration was conducted as a routine activity in the mid-term examinations of the first term in the academic year 2005-2006. Thus, the test administration was a formal activity as an academic agenda held by Diploma III Program in Tourism, Merdeka University Malang. As such, all students were aware that the test was aimed at examining their learning progress that later will constitute a part of their final score in their academic report.

The test was to be performed in Room 3,4 and 5 at the campus of Diploma III Program in Tourism, Merdeka University Malang, 1 Bandung Street, Malang. In addition, the test was to be scheduled for 90 (ninety) minutes, starting from 8.00 to 9.30 on Wednesday, 23 November 2005. As a formal examination activity, the students' seating was arranged in such a way that everyone was seated on every other classroom chair. Two proctors were assigned for the purpose. This was a standard conduct exercised in Diploma III Program in Tourism, Merdeka University Malang. First, the students did the test. When the time was called for them to stop working, they were then given an evaluation sheet to complete (see Appendix 5). The purpose was to examine their personal responses about the test they had just had.

##### **4.1.2 Test Takers**

In all there were 142 (one hundred and forty two) students taking part in the mid term examinations. They were undergraduate students of Hotel Management Department, Diploma III Program in Tourism, Merdeka University Malang in their fifth semester of their study time.

## 4.2 Analysis of Tryout Test Results

Prior to data analyses, several conducts were to be performed. The purpose was to prepare and manage the data collected so as to ensure that the analysis could be carried out with non technical problems. The activities include: (1) checking the completeness of the students' answer sheets in terms of the number as well as the number of the pages in each individual's student answer sheet; (2) putting a code number on each student's answer sheet as: 1, 2, 3, 4, 5 etc. as many as the number of the students as the students' identity needed for further analyses; (3) scoring the students' performance on the test by basing on Key Answer Guide and The Scoring Scheme established; (4) recording the students' scores and their grades in Score Recording Table and Students' Data Recording Table respectively (see Appendices 10 and 11); (5) putting in the data into computer programs relevant to the types of analysis required.

The purpose of analyzing the scores obtainable from the test after it was administered was first to describe the test takers' achievement in their reading skills in the course Professional English based on the grading criteria established. Then, the analysis was also aimed at describing the characteristics of the test in terms of its item properties, such as level of difficulty and discrimination power as well as test reliability and test validity. The analysis was performed by using a computer program called *SPSS Version 10.5 for Windows* and some other data were analyzed manually by using a scientific calculator *Casio fx-4200P*. The following is the description of the results of the analyses carried out according to the types of analysis required.

### 4.2.1 Description of Students' Achievement

The complete results of descriptive analytical analyses on the students' achievement in the course are presented in Appendix 14. The students' achievement is seen from their raw score and their grade. As has been described in Chapter III, students' raw score is determined on the basis of dichotomous scoring, in which 1 (one) point is assigned to a correct answer; 0 (zero) to an incorrect answer to each item; meanwhile, students' grade is obtainable from the differing weightings applied to the scores derived from particular items. The table that follows summarizes several points of the students' data analyzed from their total raw scores and total grades achieved.

Table 4.1: Summary of Statistical Descriptions of Students' Achievement

Description	Total Raw Score	Total Grade
N Valid	142	142
N Missing	0	0
Mean	6.0845	15.8076
Std. Deviation	3.2779	10.0987
Variance	10.7446	101.9835
Range	18.00	58.33
Minimum Achieved	1.00	1.19
Maximum Achieved	19.00	59.52
Possible Minimum	0.00	0.00
Ideal Maximum	31	84

As shown in the table, the mean scores of the students seen both from their total raw score and total grade indicate that on the average the students' achievement is considered far much lower (means = 6.0845 of the total raw score and 15.8076 of the total grade) than the values of the maximum ideal, being 31 points in the total raw score and 84 points in the total grade.

In addition to this, their scores are also spread as shown from the values of standard deviations (3.2779 and 10.0987), variances (10.7446 and 101.9835) as well as ranges (18,00 and 58.33) that are considered to be sizeable. These values can mean that the variation of the learning achievement that the students scored is large. Also, this means that the students' achievement is distributed largely heterogeneously.

Seen from both the maximum scores and grades, being 31 points and 84 points respectively, that the students managed to accomplish, it was also obvious that their attainment in the learning is considered below the learning target. They only achieved 19 points in their raw score and 59.52 points in their grade at maximum. Although no one scored the minimum, their achievable minimum score and grade (1.00 and 1.19 respectively of the maximum score and grade suggest that the students' learning achievement in the course is low, which confirms their achievement in the maximum score and grade.

The facts shown in Table 4.1 above are further confirmed from the results of analyses that are focused on the frequency of their grade achievement. The summary of this frequency is presented in Table 4.2. As shown in the table, the variation of the grades is large. This is

suggested by occurrences of the grades in the figures spreading from 1.19 points to 59.52 points in the table. Furthermore, it can be seen in the table that only two students scored the highest, being 59.52 points.

Table 4.2: Frequency of Students' Grade Occurrences

Grade	Frequency	Percent	Valid Percent	Cumulative Percent
1.19	8	5.6	5.6	5.6
2.38	2	1.4	1.4	7.0
4.76	6	4.2	4.2	11.3
5.95	8	5.6	5.6	16.9
7.14	14	9.9	9.9	26.8
7.76	2	1.4	1.4	28.2
8.33	1	.7	.7	28.9
9.52	10	7.0	7.0	35.9
10.71	7	4.9	4.9	40.8
11.90	2	1.4	1.4	42.3
13.09	6	4.2	4.2	46.5
14.28	2	1.4	1.4	47.9
15.47	7	4.9	4.9	52.8
16.66	2	1.4	1.4	54.2
17.85	9	6.3	6.3	60.6
19.04	9	6.3	6.3	66.9
20.23	6	4.2	4.2	71.1
20.24	2	1.4	1.4	72.5
21.42	4	2.8	2.8	75.4
22.61	6	4.2	4.2	79.6
22.62	2	1.4	1.4	81.0
23.80	4	2.8	2.8	83.8
25.00	3	2.1	2.1	85.9
26.19	4	2.8	2.8	88.7
27.38	4	2.8	2.8	91.5
28.57	2	1.4	1.4	93.0
30.95	2	1.4	1.4	94.4
32.14	2	1.4	1.4	95.8
33.33	2	1.4	1.4	97.2
36.90	2	1.4	1.4	98.6
59.52	2	1.4	1.4	100.0
Total	142	100.0	100.0	

If the students' achievement described in Table 4.2 above is compared to the *Pedoman Penyelenggaraan Pendidikan Tahun Akademik 2004/2005, Universitas Merdeka Malang*, Table 3.6: Score Transformation, it is quite evident that only two students get grade C (56-65) since their achievement figures at 59.52 points. As a consequence, the rest get grade E (0-44) for the highest grade score after 59.52 is 36.90, which is obtained by two students. With this, then most students have failed to achieve the pre-stated objectives in the course.

The low achievement in the students' learning turns out not to be worsened by their evaluation on the test. The results of the analysis on the students' personal responses about the test as summarized in Table 4.3 show clearly a tendency of positive standings on the part of the students as seen from several aspects of the test.

Table 4.3: Students' Evaluation on the Test Tryout

Aspects	Opinion Choices	Frequency	Percentage
Instruction	1. not clear	3	2.08%
	2. fair	136	95.77%
	3. very clear	5	2.15%
Wording	1. not clear	2	1.40%
	2. fair	134	94.36%
	3. clear	6	4.24%
Difficulty	1. easy	0	0%
	2. moderate	120	84.51%
	3. difficult	22	15.49%
Time Allotment	1. short	128	90.14%
	2. right	14	9.85%
	3. long	1	0.01%
Layout	1. bad	2	1.40%
	2. fair	133	93.66%
	3. good	7	4.93%
Gradation	1. bad	3	2.11%
	2. fair	130	91.54%
	3. good	4	6.34%
Effect	1. distressing	9	6.33%
	2. so and so	91	64.08%
	3. challenging	42	29.58%

If the students were genuine enough in giving their responses, they seemed to have considered the test 'fair' in most aspects of the test (> 64% in 'fair category'). Even, almost one-thirds (29.58%) considered the test to be challenging. The time allotment factor,



however, remained a constraint to most of the students (90.14%). The students' personal responses just described seemed similar to the respondents' evaluation on the test during informal trial of the test (see Chapter III, section 3.10.2, Table 3.13: Post-Informal Trial Data Recapitulation). In short, it might be concluded that students were positive in their evaluation in spite of the fact that they scored low in their learning achievement. Or, it may be interpreted that the students admitted indirectly that they needed to study harder in order to cope with the challenging test tasks in order to score better. This argument might be reasonable considering the fact that in order for the students of Diploma III in Tourism to perform linguistically well in their future professions, or in on-the-job training, a solid command of English is unavoidable.

#### 4.2.2 Item Analysis

To examine the characteristics of the test, item analysis was carried out (Heaton, 1988:178-184). Considering the test format employed in the present test, that is a short answer type, two test parameters were examined. These are level of difficulty of the test or facility value (FV) and discrimination. Unlike the multiple-choice type, which has a computer program to analyze, that is *ITEMAN* (Item Analysis), the short answer type provides a rare computer program in the market. Therefore, it should be reported here that the analysis to arrive at the output of the test characteristics - level of difficulty of the test or facility value (FV) and discrimination - was performed manually.

The data used to analyze these two test characteristics were the students' raw scores in which dichotomous scoring was used. This strategy was adopted because examination into these test characteristics requires a computation of the proportion of the correct and incorrect responses, which can be satisfactorily met by dichotomous scoring (Linn and Gronlund, undated 315-328).

For the purpose of analyzing the test items, several steps recommended by Heaton (1988:174-184) were performed as follows:

- 1) Putting the students' raw scores in an ascending order according to the rank;
- 2) Taking out one-thirds of the students from the top and the bottom rank, yielding about 48 students in the upper group (U) and another 48 students in the lower group (L);
- 3) Counting the number of the students who scored each item in the test correctly in both upper (U) and lower (L) groups;

- 4) Applying the formula:  $P = R/T \times 100$  for item difficulty and  $D = R_u - R_l / .5T$  for item discriminating power (Linn and Gronlund, undated 320-321), where  $P$  = index of item difficulty,  $R$  = the number of the students who got the item right:  $T$  the total number of the students who tried the item (both in the upper and the lower groups);  $D$  = index of discrimination power,  $R_u$  = the number of the students who got the item right in the upper group,  $R_l$  = the number of the students who got the item right in the lower group;
- 5) Evaluating the index obtained by comparing its value to the criteria as follows:

Table 4.4: Index and Category of Item Difficulty

Proportion Correct ( $P$ )	Category
$P > .70$	Easy
$.30 \leq P \leq .70$	Moderate
$P < .30$	Difficult

(Pusat Penelitian dan Sistem Pengujian, 1997:18)

Table 4.5: Index and Category of Item Discriminating Power

Proportion Correct ( $P$ )	Category
$D > .40$	Satisfactory
$.30 \leq D \leq .40$	Good (with a slight revision)
$.20 \leq D \leq .29$	Marginal (needing revision)
$D < .19$	Bad (omitted or total revision)

(Naga. 1992:69)

The complete results of the analysis of these two item characteristics are presented in Appendix 12. Table 4.6 summarizes the results of the analysis. As shown in the table, it is evident that the test items are mostly difficult to students (25 items or 81%). In terms of discrimination power, the items are almost evenly distributed. There are slightly more marginal and bad items with the same number (11 items each or 34.38%) than good items (6 items or 19%) and very good ones (3 items or 9.38%) .

Table 4.6: Summary of Results of Item Analysis

Item	<i>P</i>	Category	<i>D</i>	Category
1	.51	moderate	.28	marginal
2	.17	difficult	.28	marginal
3	.46	moderate	.44	very good
4	.23	difficult	.21	marginal
5	.32	moderate	.39	good
6	.61	moderate	.39	good
7	.47	moderate	.36	good
8	.43	moderate	.44	very good
9	.30	difficult	.39	good
10	.27	difficult	.28	marginal
11	.21	difficult	.42	very good
12	.17	difficult	.34	good
13	.13	difficult	.26	marginal
14	.13	difficult	.21	marginal
15	.10	difficult	.21	marginal
16	.13	difficult	.26	marginal
17	.23	difficult	.26	marginal
18	.10	difficult	.21	marginal
19	.17	difficult	.34	good
20	.09	difficult	.18	bad
21	.10	difficult	.10	bad
22	.28	difficult	.26	marginal
23	.06	difficult	.07	bad
24	.10	difficult	.10	bad
25	.01	difficult	.02	bad
26	.00	difficult	.00	bad
27	.07	difficult	.05	bad
28	.11	difficult	.13	bad
29	.07	difficult	.15	bad
30	.09	difficult	-.07	bad
31	.03	difficult	.05	bad

#### 4.2.3 Evidence of Test Reliability

Test reliability is estimated by using coefficient Alpha, which is close to the KR-20 coefficient. For the purpose, therefore, the students' data in the grade values were used because these data were close to continuous scoring, which is required of the Alpha coefficient to be run, in contrast to dichotomous data in the analysis of the KR-20 coefficient (Cohen, et. al, 1988:111). The following formula is employed:

$$r_a = \left( \frac{k}{k-1} \right) \left( 1 - \frac{\sum \sigma_i^2}{\sigma^2} \right)$$

30 where:  $r_a$  = coefficient Alpha,  $k$  = the number of items,  $\sigma_i^2$  = the variance of one item,  $\sum \sigma_i^2$  = the sum of variances of each item, and  $\sigma^2$  = the variance of the total test scores. The results of the analysis of the computation using the *SPSS 10.5* package for *Windows* are presented in as follows:

Table 4.7: Summary of Results of Reliability Analysis

RELIABILITY ANALYSIS - SCALE (ALPHA)					
*** ITEM26 46 has zero variance					
N of Cases = 136.0					
Item Variances	Mean	Minimum	Maximum	Range	Max/Min
Variance	.9441	.0425	5.7516	5.7092	135.3846
1.1453					
97 Reliability Coefficients 30 items					
Alpha = .6683 Standardized item alpha = .6920					

As shown in the table above, item 26 was excluded from the analysis because it has a 0 (zero) variance. Also it is evident that the reliability coefficient using coefficient Alpha figures at .6683 with its standardized item alpha equaling to .6920 or, if rounded up, .70. These figures indicate that, for classroom purposes, (cf. Nunnally, 1981:245) the test reliability is considered sufficient, which can be interpreted that the 30 (thirty) test items as a whole, exempting item 26, are, it is argued, internally consistent and homogeneous in measuring a single trait, that is reading ability as is aspired in basic competence in the course. This 'allows relatively straightforward test-score interpretation' (Cohen, et. al, 1988:109), in

that inference '...from the responses to a specific set of items the degree to which a [student<sup>1</sup>] has mastered the essential skills...' (Mehrens and Lehman, 1982:285) in this case reading skills is securely stable.

#### 4.2.4 Test Validity

The validity of the test in this report is viewed from several aspects. These are evidence of content validity, construct validity, empirical validity, item validity, face validity and washback validity. The following is the description of the report in each of these.

##### 4.2.4.1 Evidence of Content Validity

A language test is said to be content valid if the test satisfies the condition that '...its content constitutes a representative sample of the language skills, structures, etc, with which it is meant to be concerned' (Hughes, 2003:27). The term 'representative' in the definition seems to be the keyword in content validity.

Coverage or Representativeness implies a universe from which a sample really can reflect the characteristics of the population represented to by the sample. This is essentially concerned with a concise agreement between the ability to be measured and the content of the test to measure the ability (Djiwandono, 1996:92). In the present test, the 'population' in this perspective would be the course objective, which is 'the ability to read English texts of the intermediate to upper intermediate levels, related to tourism industries, particularly hotel-management matters'. This objective is general in nature and is not readily measurable as such. For measurement purposes, there is a highly technical requirement that the objective be derived into its more operational terms (Bachman, 1990:244). In the present test, the course objective being the general domain has been derived in terms its 'micro skills', which are further analyzed into their corresponding indicators. As has been described in Chapter III, the resulting indicators as determined in the reading course in the present report are as follows:

1. locating factual information
  - finding main factual information: *who, what, where, and/or when*;
  - finding factual explanatory: *why* and *how*
2. determining propositional inferences of a text
  - finding propositional informative inferences
  - finding propositional exploratory inferences

---

<sup>1</sup> added

3. evaluating expressions in a text
  - identifying values of expressions in the forms of facts, opinions, evidence, definition, or implication
4. summarizing and inferring underlying information in a text
  - drawing a logical conclusion,
  - predicting outcomes, and/or
  - generalizing
5. understanding information in a text through visual presentations
  - presenting information contained in a text through visual presentations like a picture, a chart, or a curve

A further step taken was to assign the items that are designed to measure these performance indicators of the ability. As to the best knowledge of the present test developer there has been no authoritative guidance as how many items are representative enough to measure indicators, the decision to involve a particular number of items is motivated by the logical judgment. The key principle is that as far as possible, a performance indicator is measured by at least two items. The consideration is that if one's reading skill is not measurable by an item, another item parallel to it is expected to secure the situation.

To meet the requirements of content validity, then, a table of test content specification is further established (see Table 3.1: Test Content Specification, p.16, Ch.3). With this test blueprint at hand, the <sup>96</sup> content relevance and the content coverage (Bachman, 1990:244; Messick, 1989) of the test being made are met. In short, it can be stated that by design the test content is securely satisfied in terms of their relevance and coverage. Thus, the content validity of the test has technically been established.

#### 4.2.4.2 Evidence of Construct Validity

Technically evidence concerning <sup>95</sup> construct validity of a test can be assessed through 2 (two) interrelated ways: examination on the construct establishment, investigation into its empirical supports (cf. Bachman, 1990:254-260). The former simply addresses what construct is and how it is created conceptually; the latter puts a more emphasis on empirical confirmation or exploration of the psychological existence of the construct (Cramer, 2003). While the former necessitates a synthesis of theories of 'something', the latter deals more closely with ways of knowing the psychological reality of that thing empirically.



With this framework in mind, evidence of construct validity of the present test necessitates the assessment of the synthesizing of reading ability and assessment of how it is empirically supported by the data. These two need to be attended to separately in the section that follows.

As has been described in section 3.3 Test Content Specification in Chapter III as well as the previous section, the test is developed to measure how far the students have made progress with their learning of reading ability. Thus, the concept to be revealed in the present report in terms of construct is reading ability. A synthesis of concepts concerning what constitutes reading ability has been discussed in some detail with a selected reference to those concepts forwarded by Alderson (2000), Gear and Robert (1996), Grabe (1991), Bachman (1990), and Munby (1978), resulting in the dimensions of 5 (five) functional competencies in reading or micro skills. In this test, it is conceptualized that these micro skills consist of reading skills that follow: locating factual information, determining propositional inferences of a text, evaluating expressions in a text, summarizing and inferring underlying information in a text, and understanding information in a text through visual presentations.

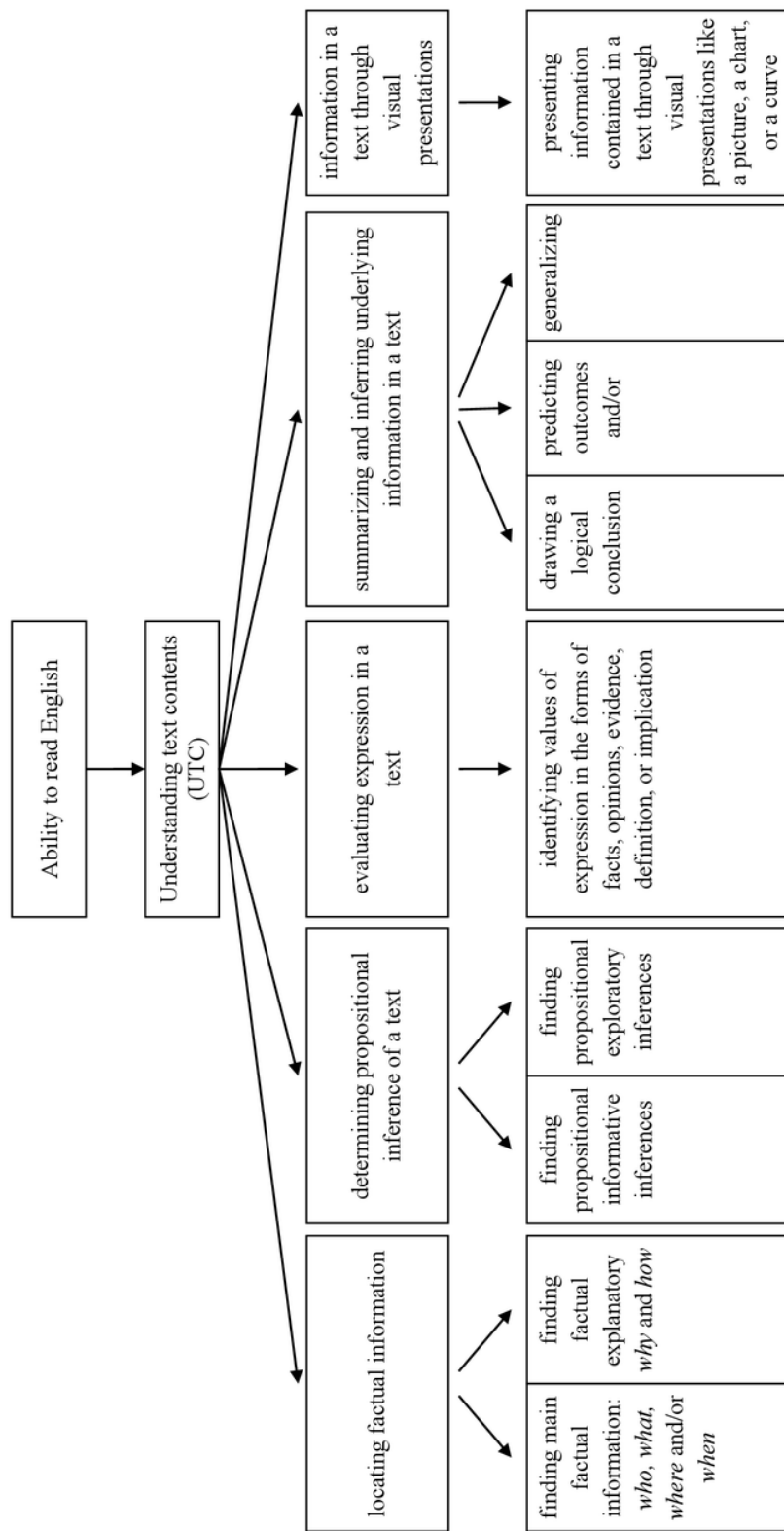
A further investigation into the works of Wiener and Bazerman (1988) and Anderson, Durston and Poole (1969) yields performance indicators of those 5 (five) functional competencies in reading or micro skills. These performance indicators that correspond to their micro reading skills are respectively as follows: finding main factual information: *who*, *what*, *where*, and/or *when*, finding factual explanatory: *why* and *how*; finding propositional informative inferences; finding propositional exploratory inferences; identifying values of expressions in the forms of facts, opinions, evidence, definition, or implication; drawing a logical conclusion; predicting outcomes, and/or; generalizing; and presenting information contained in a text through visual presentations like a picture, a chart, or a curve.

These performance indicators are then used as a basis of writing relevant test items. Up to this point then basically a process of construct establishment has been underway. Thus, in this perspective, partially the present test actually is endowed with a conceptual basis. In short, there is evidence of construct validity in the present test.

Therefore, conceptually the structure that underlies reading ability examined in the present test can be diagrammatically illustrated as follows:



Figure 4.1: The Structure of Reading Ability Conceptualized



The next task to find empirical bases of such a concept of reading ability requires field testing of the test. To field test the test developed, there is an array of approaches, starting from the simple to the sophisticated ones Bachman (1990:258). The simple one is through reliability analysis discussed previously in section 4.2.3 Test Reliability. For the purpose of the present test, coefficient Alpha is used. It turns out that the coefficient Alpha of the present test is .6920 rounded up to .70, which is interpreted that, exempting item 26, the test items are internally consistent and homogeneous in measuring a single trait, that is reading ability as is aspired in basic competence in the course.

Of the empirical procedures, A possible approach that can be taken to examine construct validity of the present test is through factor analysis. Also, there is another complicated procedure called the multi method-multi trait approach (Bachman, 1990:263). The latter, however, is beyond the reach of the present test because it requires the use of differing means of assessing the trait under interest. For the purpose of evaluation of evidence of construct validity of the present test, besides coefficient Alpha, factor analysis is attempted, certainly with some technical constraints that exist such as limitation of the number of the subjects taking part in the validation.

The analytical procedures use a computer program called *SPSS Version 10.5 for Windows* based on a step-by-step data analysis outlined by Norusis (1993:47-81) with a conceptual view of Cohen et.al's account on factor analysis (Cohen, et. al, 1988:621- 628). The complete results of the analysis are presented in Appendix 15. The following is the descriptive summary of the results, presented in the order of the analysis.

The first step to perform was to examine the variables that are adequate statistically to be analyzed. The term 'variable' is used as a statistical term. In this context these variables under interest are actually the items of the test<sup>2</sup>. The first attempt made then is examination of all the 31 (thirty one) items of the test. In the first analytical attempt, it turns out that there was a warning in the output of the analysis on the computer that the variables could not be analyzed further because one of the variables had a 0 (zero) variance. When examined further, the variable with this 0 (zero) variance was variable 26 or item 26. This variable was then excluded from the analysis, leaving 30 (thirty) variables for a further analysis.

---

<sup>2</sup> In the discussion of factor analysis results the term variables is used to mean items.

Reanalysis on the 30 remaining variables is still meant to examine the adequate variables. This procedure is actually a process of satisfying the requirement for factor analysis to proceed. The indicator of fulfillment of the requirement is shown by a measure called Kaiser-Meyer-Olkin Measure of Sampling Adequacy (or KMO). A minimum index required is .50. This process is iterating in the sense that data reanalysis can be performed again and again by reducing the variables that are not statistically adequate to be analyzed. The criterion variable exclusion is the same as the value of the KMO index, that is, minimum .50.

The process of examining variable adequacy took 13 (thirteen) iteration cycles, that is, when the KMO measure exceeds .50 with all variables shown in the anti-image matrices in the analysis output demonstrating a correlation coefficient greater than the minimum required, being .50<sup>a</sup>. This process resulted in a value of the KMO measure figuring at .698 (sig. at the level of .000) (See Appendix 15) and the exclusion of several inadequate variables which are variables (presented in the order of iteration) 10, 22, 21, 4, 25, 3, 5, 30, 31, 24, 9, 7, 23 and 8, thus leaving adequate items with their indices of correlation 1 (.582<sup>a</sup>), 2 (.755<sup>a</sup>), 6 (.570<sup>a</sup>), 11 (.633<sup>a</sup>), 12 (.801<sup>a</sup>), 13 (.799<sup>a</sup>), 14(.716<sup>a</sup>), 15 (.680<sup>a</sup>), 16 (.851<sup>a</sup>), 17 (.642<sup>a</sup>), 18 (.569<sup>a</sup>), 19(.603<sup>a</sup>),20 (.681<sup>a</sup>), 27 (.763<sup>a</sup>),28 (.597<sup>a</sup>), and 29 (.674<sup>a</sup>) (See Appendix 15 shown in anti-image correlation section in anti-image matrices).

The next phase to be performed was to further analyze the variables considered adequate statistically. This process includes 2 (two) main stages: Factoring and Rotation. Factoring is a process of creating 'new factors' for the selected adequate variables. In this process variables are put into their appropriate factors. The complete results of the analysis in this stage are presented in Appendix. Table 4.8 presents the summary of the results.

As Table 4.8 shows there are 5 new factors or components created with their members of variables. Most variables indicate significant values ( $\geq .50$ ); however, two indicate insignificant values ( $\leq .05$ ). The members of each factor or component are as follows: Factor 1: item 2, item 12, item 13, item 14, item 16, item 17, item 27; Factor 2: item 6, item 11, item 18; Factor 3: item 19 and 20; Factor 4: item 28, Factor 5: item 1. Item 29 belongs to none of the factors because the value is  $\leq .05$  in all factors created. All these results, however, need further confirmation to ensure that all the variables have been accurately placed into their new factors. A next stage called 'rotation' then is run the purpose of which is confirm the placement of the variables in their corresponding factors.

Table 4.8: Unrotated <sup>26</sup>Component Matrix

Item	Component/Factor				
	1	2	3	4	5
<sup>10</sup> Item 1	-	-	-	-	.607
Item 2	.613	-	-	-	-
Item 6	-	.583	-	-	-
Item 11	-	.562	-	-	-
Item 12	.614	-	-	-	-
Item 13	.608	-	-	-	-
Item 14	.669	-	-	-	-
Item 15	-	-	-	-	-
Item 16	.650	-	-	-	-
Item 17	.532	-	-	-	-
Item 18	-	-.559	-	-	-
Item 19	-	-	.683	-	-
Item 20	-	-	.509	-	-
Item 27	.664	-	-	-	-
Item 28	-	-	-	-.579	-
Item 29	-	-	-	-	-

The Complete results of rotation are presented in Appendix 15. Table 4.9 summarises results of rotating the variables.

26  
Table 4.9: Rotated Component Matrix

Item	Component/Factor				
	1	2	3	4	5
10 Item 1	-	-	-	.767	-
Item 2	.651	-	-	-	-
Item 6	-	.620	-	-	-
Item 11	-	.743	-	-	-
Item 12	-	.675	-	-	-
Item 13	.786	-	-	-	-
Item 14	.549	-	-	-	-
Item 15	-	.632	-	-	-
Item 16	-	-	-	-	.503
Item 17	-	-	-	-	.533
Item 18	-	-	-	.763	-
Item 19	-	-	.855	-	-
Item 20	-	-	.619	-	-
Item 27	.632	-	-	-	-
Item 28	-	-	-	-	.719
Item 29	-	-	(.487) ?	-	-

As shown in Table 4.9, the values of the variables are enhanced, thus indicating a firmer of placement or displacement of variables. Also it is evident that variables formerly placed under a certain factor as shown in Table 4.8 are placed in their new components. Factor 1 has 4 (four) members: items 2, 13, 14, and 27; Factor 2 holds items 6, 11, 12, 15; Factor 3 keeps items 19 and 20; Factor 4 has items 1 and 18, and finally Factor 5 maintains items 16, 17 and 28. What is further confirmed is that in spite of its significant value as shown by its adequacy, item 29, by criterion of significance minimum .50, belongs to none of the factor created. The possible placement of this variable is putting it into Factor 3 based on the value figuring at .487 which is closer to .50 as the minimum criterion.

Rotation can be considered a final step in factor analysis if the purpose is to examine the adequate variables with their placement. There are other optional steps that may be carried out: factor validation and factor scoring. In the present analysis, these two were not performed because the purpose of the present analysis has been achieved, that is to examine empirically adequate items of the test.

Based on factor analysis described previously, it is evident that the present test has construct validity. It is shown first that there are empirically 5 (five) factors similar to the 5 (five) functional competencies in reading or micro skills conceptualized for the present test. In addition, there are items that can adequately be used to measure these 5 (five) functional competencies in reading or micro skills. In other words, the test items that are assigned adequately to measure these skills are empirically established. Based on this then, to measure reading functional or micro skills, the present test consists of these empirically-supported items.

#### 4.2.4.3 Evidence of Empirical Validity

Empirical validity is another term used for criterion-related validity. It is 'a judgment regarding how adequately a test score can be used to infer an individual's most probable standing on some measure of interest – the measure of interest being the criterion' (Cohen et. al, 1988:128). Of the two types of empirical validity – predictive and concurrent – the latter was examined in the present test. The rationale is that of the data available needed for the purpose of analyzing empirical validity, data on the students' future performance is not available yet. The accessible data available relevant for the purpose of empirical validity analysis was students' grade point average (GPA). Therefore, the type of validity examined in the present test is concurrent validity. In this context, then the students' grade point average (GPA) is put to stand as the external criterion. In addition to this, two types of scores were included in the analysis: raw scores and grades. The purpose is to examine the consistency of closeness between the test scores of the test and the external criterion as well as the consistency of scores within different scoring procedures.

The analysis of the data for the purpose of empirical validity examination is processed by using a computer program called *SPSS Version 10.5 for Windows*. The statistical procedure used is the formula for the Pearson bivariate product-moment correlation coefficient. According to Brown (1996:155), the formula for calculating the coefficient is expressed as follows:



$$r_{xy} = \frac{\Sigma(Y - \bar{Y})(X - \bar{X})}{NS_y S_x}$$

where

$r_{xy}$  = Pearson product-moment correlation coefficient

$Y$  = each student's score on Test Y

$\bar{Y}$  = mean of Test Y

$S_y$  = standard deviation on Test Y

$X$  = each student's score on Test X

$\bar{X}$  = mean of Test X

$S_x$  = standard deviation on Test X

$N$  = the number of students who took the two tests.

The results of the analysis is presented in Table 4.10. As the table shows, the students' raw scores and their grade positively and significantly correlate with their grade point average (GPA) at the level of  $\alpha = 0.05$  (2- tailed) figuring at  $r_{xy} = .179$  and  $r_{xy} = .185$  respectively. Although the closeness of the scores is weak (Linn and Gronlund, undated 61-62)<sup>3</sup> as shown by their values, this indicates that scores obtainable by using the reading test can be used, with weak accuracy to indicate the students' GPA achievement.

Table 4.10: Summary of Concurrent Validity Analysis

		GPA	RAW SCORE	GRADE
GPA	Pearson Correlation	1.000	.179*	.185*
	Sig. (2-tailed)	.	.033	.027
	N	142	142	142
RAW SCORE	Pearson Correlation	.179*	1.000	.893**
	Sig. (2-tailed)	.033	.	.000
	N	142	142	142
GRADE	Pearson Correlation	.185*	.893**	1.000
	Sig. (2-tailed)	.027	.000	.
	N	142	142	142

\* Correlation is significant at the 0.05 level (2-tailed)

<sup>3</sup> The relationship between the raw scores and the grades is expected to be strong as it is shown in the results of analysis being  $r_{xy} = .893$  significant at the 0.01 level (2-tailed) because these scores are essentially identical.



<sup>33</sup>  
\*\* Correlation is significant at the 0.01 level (2-tailed)

#### 4.2.4.4 Evidence of Item Validity

Another important feature of a test that needs to be examined deals with item validity. Item validity actually belongs to the discussion of reliability because fundamentally it deals with internal consistency (Gall and Borg, 1983:284-285) among items from a single test administration (Linn and Gronlund, undated:89). For the purpose, a procedure known as <sup>90</sup>Kuder-Richardson formula 20 or KR-20 is used. The procedure is used for scores that are scored dichotomously (1 or 0) in which the computation involves the examination of the proportion of persons answering each item correctly and the standard deviation of the total scores (Djiwandono, 1996:152; Nitko, 1996:69)

Internal consistency is expressed by using a coefficient of internal consistency. The computation of the coefficient involves the following formula (adapted after Djiwandono, 1996:152):

$$r = [k : (k-1)] \times [1 - (\sum pq : S^2)]$$

where:

$r$  = coefficient of internal consistency

$k$  = the number of items

$p$  = proportion of students' correct answers

$q$  = proportion of students' incorrect answers

$S^2$  = variance of total scores.

In this present report, the computation was performed by using a computer program called *SPSS Version 10.5 for Windows*. The entire <sup>26</sup>results of the analysis can be seen in <sup>26</sup>Appendix 16. The summary of the results is presented in Table 4.11. The table presents the coefficient of internal consistency all the 31 (thirty one) items in the test. It is evident from the table that the items have different values of the internal consistency coefficient with or without different level of significance.

Based on the level of consistency of 0.05 and/or 0.01, most items (77.41%) are categorized as empirically valid items. These are evidenced from items 1, 2, 3, 4, 5, 6, 7, 8, 9,

11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 22, 24, 27, 28, and 29. The rests, items 10, 21, 23, 25, 25,30, and 31, are not empirically valid.

Table 4.11: Summary of Results of Item Analysis

Item	<i>r</i>	Category
1	.241**	valid
2	.411**	valid
3	.290**	valid
4	.196*	valid
5	.347**	valid
6	.398**	valid
7	.331**	valid
8	.406**	valid
9	.270**	valid
10	.155	not valid
11	.442**	valid
12	.455**	valid
13	.473**	valid
14	.515**	valid
15	.386**	valid
16	.530**	valid
17	.305**	valid
18	.253**	valid
19	.356**	valid
20	.332**	valid
21	.107	not valid
22	.168*	valid
23	.113	not valid
24	.312**	valid
25	.059	not valid
26	@	not valid
27	.338**	valid
28	.333**	valid
29	.313**	valid
30	-.028	not valid
31	.100	not valid

\* Significant at the 0.05 level (2-tailed)

\*\* Significant at the 0.01 level (2-tailed)

@ Cannot be computed because at least one of the variables in constant

#### 4.2.4.5 Evidence of Face Validity

Face validity of the present test is not examined like other types of validity because to the best knowledge of the present test developer there has been no literature that deals with the estimation of face validity using statistical computations. In spite of absence in its scientific concept, it is important (Hughes, 1989:27), Face validity can be defined as '... the extent to which the test appeals to test takers and test users' (Bachman and Palmer, 1996:42). Grondlund (1998) adds that a test has face validity when the test takers consider it '... as fair, relevant, and useful for improving learning.' A test that has face validity is the one that '... looks as if it measures what it is supposed to measure' (Hughes, 1989:27). The present test is a test of reading comprehension. With this concept then the reading test is the test that measures students' reading comprehension. In this present test, the evidence of the test face validity is inferred indirectly from two sources. These are the moderators and the students as the test takers.

Seen from the moderators' response, particularly the one in the aspects 'agreement' and 'function' in the feedback evaluation sheet, it is obvious that of the 3 (three) moderators, no one provided a response of 'unrelated' when they were to rate the agreement of whether the items referred to in the test measure their corresponding performance of indicators of reading. Even, the expert moderator suggested that items that measure other than understanding of reading text content be excluded. Similarly, in the aspect 'function' which essentially asks whether the test can be appropriately functioned as an achievement reading test, no one of the moderators responded 'bad'.

Evidence of face validity from the students' response is seen from their consistent feedback both during the informal trial and the test tryout. Most (66.67% out of 6 in the informal trial stage and 29.58% out of 142 in the tryout) expressed their feeling that they were challenged to do the test. In addition, the students' worksheet can also be used as evidence of face validity of the test. Apart from their correctness or incorrectness of their response, they certainly needed to read all the texts in the test as instructed before they really responded to all the test tasks. In short, it can be stated that the test has face validity. The test is a direct, performance test in that the students actually perform reading in order to respond to test tasks.

#### 4.2.4.7 Evidence of Washback Validity

Heaton (1988:170) once mentioned backwash effect of a test on teaching. Brown (2004:26) mentions sequential validity. Hughes (2003:57) prefers backwash. When Heaton's, Brown's and Hughes's ideas are compared, they basically can be traced back to the same root: what effect (s) the test impacts. When Heaton focuses on teachers' teaching, and Hughes (2003:53-57) extends it to reach students' learning, Brown includes a broader scope to include <sup>88</sup> not only technical aspects of language testing but also a social spectrum.

Washback or backwash effects of the present test has not been able to be reported here because the effects of the present test need not only time for the impact to process after the administration of the test but also a careful design to carry out, which is still beyond the scope of the present test development. However, a slight hint of impact that may have arisen is seen from the students' response when they state that the test challenge them.

#### 4.2.5 Evidence of Practicality

Practicality of a test is also another important facet of a test. <sup>87</sup> A test that is highly reliable and valid in its all dimensions yet impractical may turn out to be not a good test (Davies, 1990). Brown (2004:19) argues that practicality <sup>61</sup> of a test is closely related to effectiveness <sup>86</sup> of a test in that a practical test is a test that does not require a large amount of financial resources. Also, a practical test is the one that remains there fitting in time constriction. Besides, it is practical in terms of ease of administration. Finally, a practical test is the one that encloses a specific and time efficient scoring scheme.

If projected to these conditions, the present test seems to bear some positive points in some respects and negative ones in others. First, although it is a relative measure, financial spending required of the test to develop is rather costly. The final version of the test requires 8 (eight) pages in all: 5 (five) pages for the test tasks and another 3 (three) for the answer sheet. For three parallel classes of 142 (one-hundred and forty two) students, there would be 8 pages x 142 students, which is equal to 1136 pieces of paper. For test multiplication by photocopying, it approximately costs about 1136 pieces x Rp. 100,- = Rp. 113.600,- (one-hundred and thirteen thousand six hundred rupiahs).

In terms of time spending and administration, the test is practical, however. By just administering the test tasks together with the answer sheet to 142 (one hundred and forty two) students in 90 (ninety) minutes, the test developer is able to examine a wide range of students' progress in their learning reading skills -- the test consists of 5 (five) performance indicators

that reflect reading abilities. In comparison with an ethnographic approach to examine students' progress in their reading skills, which requires a close-encounter with all 142 (one hundred and forty two), the test is certainly more practical.

The test is also considered practical with its scoring scheme. Although it requires some practice for familiarity to use the scoring procedure, objectivity as well as accuracy in scoring is securely maintained. In short, it seems that the positive sides of the test outweigh the negative ones. Therefore, it is logical to consider the test practical.

## CHAPTER V CONCLUSION AND SUGGESTION

### 5.1 Test Development in Retrospect: Conclusions

<sup>85</sup> Based on the conduct of reading comprehension test tryout described in the previous chapters, several important points can be drawn as conclusions as follows. First, when carried out based on the true principles of language testing and assessment, language test development is a sophisticated undertaking that requires critical and careful thoughts. For example, 'translating' performance indicators into items that are expected to measure the indicators necessitates considerations of several aspects: target audience, level difficulty of the expected item, test format, accuracy of wording, ease of scoring, authenticity of test tasks as well as test administration. Realized or not, a novice test developer, particularly those like the present test writer, uses his/her personal judgment much more than these principles. As such there is a tendency to employ non technical expertise during the test development practices, which is certainly not desirable by concept but it really tends to happen.

Next, if principles of language testing have as far as possible been tried to be satisfied during test development conceptually, what is thought to have been there with these principles in the test items, it can turn out that things can go differently in practice. For example, an item considered to be difficult by design may come up statistically to be easy in practice, or vice versa. Another case is that an item is supposed to measure a particular performance indicator by design, however, when analyzed statistically, it measures another distinct performance indicator. In such a case it seems that the test developer's accuracy and sensitiveness to the characteristics of the test taker plays an important role as an initial 'needs assessment' on which to be further principled judgement to develop language tests.

A good test is the one that by concept can fulfill all important facets of test requirements: quality of items, reliability, validity and practicality a professional English teacher inevitably must address in testing practices. Developing a language test is then a challenging business, which it seems that many take it for granted. The tryout just reported is a good exercise to reflect the real picture of how fragile it is to manage the fulfillment of requirements of a good test. Many important aspects of language test are still there, challenging to be addressed for improvement.

## **5.2 Suggestion**

To classroom teachers, testing communicative language performance, in spite of its exacting requirements, cannot be taken for granted as such if accountability becomes the teacher's concern. The wrong decision that a teacher makes based on the evidence as shown by scores or grades of language tests may potentially change one's fate. Therefore, a teacher needs to be open to students in linking testing practices and teaching conducts. For example, there seems to be once again a desirable practice where introduction of teaching objectives, ways how to achieve them, and accurate indicators of achievement prior to the teaching learning process are emphasized. With this in mind, students will have a clear picture of the course that will be useful for them to manage their learning strategies to meet the requirements of the course.

Practice makes perfect and practice enhances learning better than knowing does. Developing language tests needs practice not just knowing. Through actual practice in making tests with various formats to assess a variety of language skills, a teacher's sensitiveness in judging one's communicative performance of English will be sharpened, which will benefit personally and professionally. Therefore, creativity in experimenting and trying to develop tests need to constitute their framework in actually making different types of test items.



## References

- Alderson, Charles. 2000. *Assessing Reading*. Cambridge: Cambridge University Press
- Anderson, Jonathan, Durston, Berry H. and Poole, Millicent E. 1969. *Efficient Reading: A Practical Guide*. Sydney: McGraw-Hill Book Company.
- Bachman, Lyle.F. 1996. *Fundamental Considerations in Language Testing*. Oxford: Oxford University Press.
- Bachman, Lyle F and Palmer, Adrian S. 2000. *Language Testing in Practice*. Oxford : Oxford University Press
- Barret, T.C. 1968. 'What is reading? Some current concepts' in H.M. Robinson (Ed.) *Innovation and Change in Reading Instruction*. The sixteenth handbook of the National Society for the Study of Education, Chicago: The University of Chicago Press.
- Brown, H. Douglas. 2001. *Teaching by Principles: An Interactive Approach to Language Pedagogy*. San Francisco: Longman
- Brown, James Dean. 1996. *Testing in Language Programs*. New York: Prentice Hall Regents
- Brown, Douglas. 2004. *Language Assessment: Principles and Classroom Practices*. New York: Longman
- Celce-Murcia. Marianne, Dornyei, Zoltan and Thurell, Sarah. 1995. 'Communicative Competence: A Pedagogical Motivated Model with Content Specifications'. *Issues in Applied linguistics*. Vol.6 No.2, pp. 5-35.
- Champeau De Lopez, Cherryl L, Marchi B, Giancarla and Arreza-Coyle. 'A Taxonomy: Evaluating Reading Comprehension in EFL'. 1997. *English Teaching Forum*, Volume 35, Number 2, April. 1997, pp. 31-32
- Cohen, Ronald J., Montague, Pamela, Nathanson, Linda S., and Swerdlink Mark E. 1988. *Psychological Testing: An Introduction to Tests and Measurement*. Mountain View: Mayfield Publishing Company
- Cramer, Duncan. 2003. *Advanced Quantitative Data Analysis*. Maidenhead, PA: Open University Press.
- Davies, Alan. 1990. *Principles of Language Testing*. Cambridge, MA: Basil Blackwell, Inc.
- Djiwandono, M Soenardi. 1996. *Tes Bahasa dalam Pengajaran*. Bandung: Penerbit ITB Bandung.
- Douglas, Dan. 2000. *Assessing Languages for Specific Purposes*. Cambridge: Cambridge University Press.

- Elijah, D., and Legenza. 1975. A comprehension taxonomy for teacher. *Reading Improvement*, 15, 1. pp. 98-99.
- Grabe W. 1991. 'Current development in second language reading research.' *TESOL Quarterly* 25, 3. pp. 375-405.
- Gray, W.S. 1960. 'The major aspects of reading' in H.M. Robinson (ed.) *Supplemnetary Educational Monographs No. 90*. Chicago: Univeristy of Chicago Press.
- Gronlund, Norman E. 1998. *Assessment of Student Achievement. 6<sup>th</sup> Edition*. Boston: Allyn and Bacon.
- Heaton, J.B. 1975. *Writing English Language Tests*. London: Longman
- Heaton, J.B. 1988. *Writing English for Language Tests*. London: Longman
- Herber, H. 1978. *Teaching Ceading in Content Areas 2<sup>nd</sup> Edition*. Englewood Cliffs: Printice-Hall
- Hughes, Arthur. 2003. *Testing for Language Teachers*. Cambridge: University Press
- Irwin, J.W. 1986. *Teaching Reading Comprehension Process*. Englewood Cliffs: Prentice-Hall Inc.
- Linn, Robert L. and Gronlund, Norman E. Undated. *Measurement and Assessment in Teaching 7<sup>th</sup> Edition..* Englewood Cliffs, New Jersey: Merrill, an imprint of Prentice Hall
- Messik, Samuel. 1989. 'Validity' in Robert L. Linn (ed.) *Educational Measurement 3<sup>rd</sup> Edition*. New York: American Council on Education, Macmillan Publishing Company. pp. 13-104.
- Naga, Dali S. 1992. *Pengantar Teori Sekor pada Pengujuran Pendidikan*. Jakarta: Penerbit Gunadarma
- Nitko, Anthony. J. 1990. *Educational Assessment of Students*. New Jersey: Prentice Hall
- Norusis, Marija J. 1993. *SPSS SPSS for Windows Professional Statistics Release 6*. Chicago, III.: Marketing Department SPSS Inc.
- Nunnally, Jum C. 1981. *Psychometric Theory*. New Delhi: Tata McGraw-Hill Publishing Company Ltd.
- Pearson, P.D. and Johnson, D.D. 1978. *Teaching Readng Comprehension*. New York: Holt Rinehart and Winston.

## Appendix 1: Reading Text Pool for Initial Text Selection

### Text 1

Most lodging and food service operations require the coordinated efforts of a large number of staff members to properly provide product and services to the guests at costs levels which generate required profits. Most staff members are part of formal work groups. When the group members work together cooperatively they become a team. A well-organized team can accomplish much more than several employees working alone or even with an uncoordinated group.

### Text 2

Groups, both formal and informal, and cliques (subgroups) are naturally formed within most hospitality operations. Effectively managed, they can be built into teams that will yield very effective group performance. Managers must know how to work with groups. This is important when supervising employees on the job and when providing training to staff members. Group training, during which the trainer serves as a facilitator (helper) of the group members' learning activities, involves application of the same basic principles as does working with groups in on-the-job performance of work tasks. Managers, then, must know how to effectively develop and work with groups when supervising and when training.

### Text 3

Nabe, needless to say, is just one of the countless Japanese foods that have not made journey overseas. Another very popular everyday food that falls into this category is Japanese noodles. As with nabe, exactly how you eat your noodles often depend upon where in Japan you eat them. Generally speaking, people in the east favor soba noodles, which are made from buckwheat flour kneaded into five brown threads, while those in the west prefer thick wheat udon noodles. Both soba and udon vary in the taste and thickness according to the region in which they are made. Most commonly served in a very simple, hot soy-sauce based broth or plain-boiled to be dipped into liquid condiment, Japanese noodles are a healthy nourishing treat. Here again, though, the foreign influence in Japanese culinary habits is very much in evidence, with curry soba and curry udon being among the most popular Japanese noodles variants.

### Text 4

The success of Karihara's book has no doubt been welcomed by the Japanese government, which he also been working to promote Japanese food overseas. Since being set up in 2004, the government's Intellectual Property Strategy Headquarters has followed a policy trying to raise the brand value and push global sales of Japanese food. Despite the good reputation of Japanese food, there are still comparatively few Japanese restaurants outside Japan. Only by increasing opportunities for people to sample genuine Japanese cooking will Japanese food ever fully establish itself as a brand in the way that Italian or Mexican or Indian food has

#### Text 5

A third group of hotels conforms to the Travelodge format. These hotels are aimed primarily at the tourist trade and business travelers using the road network. Again the market demand is such that the group operating these hotels is unable to outbid other users in the city center or at its edge, and is thus forced into edge-of-city locations, reflecting the preference of their customers.

#### Text 6

The prevailing low level of land costs results in much lower room rates, so that these hotels are attractive to a large tourist market. The importance of land costs for this sector is revealed by the fact that room rates for the Formula 1 brand are 43 a night higher in the UK than in France, a reflection of the difference in land costs between the two countries.

#### Text 7

There are so many day-to-day things that must be done at a hotel; guests must be checked-in, rooms cleaned, meals cooked, phones answered. Each activity is an important part of hotel's overriding goal of good guest service. However, the daily routine of doing the things that must get done is not the same as a hotel working toward a set of goals that everyone in a hotel understands and accepts. Ordinary hotels 'go through the motions' of performing required duties. Excellent hotels seem to be doing much more than executing the various technical steps required to house and feed guests. The management and staff of excellent hotels understand why they are doing things in a certain way as well as for whom they are doing them.

#### Text 8

During the GM research, I spent time with executives and staff of the outstanding, excellently managed hotels. It was rather remarkable to discover that the management and staff of these hotels felt that their hotel was by far the best in the city! The reasons they gave for being the best, to be sure, varied from one hotel to another. Still, this feeling of being the best was widespread among the management and staff and unmistakable in each hotel. The GMs of these hotels had developed among their management and staff a self-concept, or culture, that said a lot about how that organization felt the hotel should be run. Call it an operating philosophy, a self-concept, or a culture; whatever it is called, it's quite clear that developing the right kind of culture is an important goal for all well managed hotels.

#### Text 10

Eight of ten GMs studied feel that understanding people is one of their main strengths. Furthermore, they considered their ability to develop people of the major reasons for their success. The fifty-three hotel department heads interviewed ranked employee motivation as

their greatest challenge. Some GMs seem to be natural when it comes to people. William Scully feel that understanding and being able to deal with all kinds of people is one of his main strengths. It turns out that his subordinates feel the same way! They refer to him as fair, approachable, able to handle all kinds of people, caring, a friend, a strong leader with an easy manner, and a person who makes you want to do a good job for him. Frank Anderson is an extrovert with a different saying for every occasion. Henri Le Sassier feels that his flair and cleverness with people springs from an artistic mother.

#### Text 11

Control is the final important management function that ensures a hotel's goals and standards are met. In this regard it is important to remember that control is action oriented and also oriented toward the future. A feedback-control process is useful repetitive situations. As the results of an activity become known, they are compared to goals and standards. If the variance between results and goals is unacceptably large, management must intervene to bring the activity back into control. Feedback control can be an effective strategy when a particular business operation repeats itself and when results can be objectively measured.

#### Text 12

Control tactics can be classified according to the particular objective of the control. In the case of feed-back control systems, the object being control is the result of some activity. Another tactic relies on control of the specific actions of a hotel's employees. Finally, personnel control attempts to control employee actions by controlling their training, attitudes, and motivation.

#### Text 13

Circumstances dictate what type of control strategy is most appropriate since each has its strong and weak points. When there is considerable knowledge of the kinds of actions that lead to desirable outcomes, specific-action controls can be effective. When it is clear what kinds of outcomes are desirable, and when these outcomes can be accurately measured, results control can be used.

#### Text 14

In some circumstances it is difficult to accurately measure outcomes, and knowledge of the most desirable kinds of specific actions is low. This presents management with its most difficult control problem. In this case, personnel-control strategies must be devised so that employees, on their own, can be relied on to do what is best for the hotel. Well-managed hotels use combinations of these three different control strategies to ensure that they meet their goals and objectives.

## Appendix 2: Test Draft for Moderation

### Directions:

1. There are 9 short texts in this test. Do all the test items following these texts;
  2. Write your answer in the answer sheet provided;
  3. Give as brief answers as possible, max. two sentences;
  4. Work independently and no dictionary is allowed to be used.
- 

### **Text 1**

Hotels, both large and small, are exciting places. Many of the people who work in hotels are drawn to them because of this excitement. It is exciting, for example, to attend a gala banquet in a fine hotel. I remember attending a testimonial dinner one evening. 400 guests, all in evening attire, were having cocktails in an area adjacent to the banquet room. When it was time to be seated, the partitioned wall separating the guests from the banquet room was retracted revealing the splendidly arranged tables. The sight was so beautiful that the 400 guests spontaneously began to applaud.

1. What does the writer actually say about hotels in the text above?
2. Which sentence in the paragraph above has the writer's main idea?
3. Why does the writer use his/her personal experience in attending a gala banquet?
4. How are the first two sentences in the text above related to other sentences after these two?
5. What is the best title for the text above?

### **Text 2**

Groups of people in hospitality industries, both formal and informal, and cliques (subgroups) are naturally formed within most hospitality operations. When they are effectively managed, they can be built into teams that will yield very effective group performance. Managers must know how to work with groups. This is important when supervising employees on the job and when providing training to staff members. Group training, during which the trainer serves as a facilitator (helper) of the group members' learning activities, involves application of the same basic principles as does working with groups in on-the-job performance of work tasks. Managers, then, must know how to effectively develop and work with groups when supervising and when training.

6. Which sentence in the paragraph above has the writer's main idea?
7. What topic does the writer talk about in the paragraph above?
8. According to the writer, what responsibilities do managers need to have?
9. What is meant by basic principles in the text?
10. Provide an example of a case where a manager works together with staff members on the job!

### Text 3

Control tactics can be classified according to the particular objective of the control. In the case of feed-back control systems, the object being controlled is the result of some activity. Another tactic relies on control of the specific actions of a hotel's employees. Finally, personnel control attempts to control employee actions by controlling their training, attitudes, and motivation.

11. What does the writer say about control tactics?
12. How many types of control tactics can be used according to the text above?
13. What technique does the writer use to show the relationship between control tactics in general and the number of their types?

### Text 4

It is perhaps easiest to grasp the hierarchy of strategic decision making by taking a look at a typical organization chart for a hotel company as shown by a figure below. The figure shows such an organization chart along with the kinds of strategic decisions that are normally made at various levels within the organization.

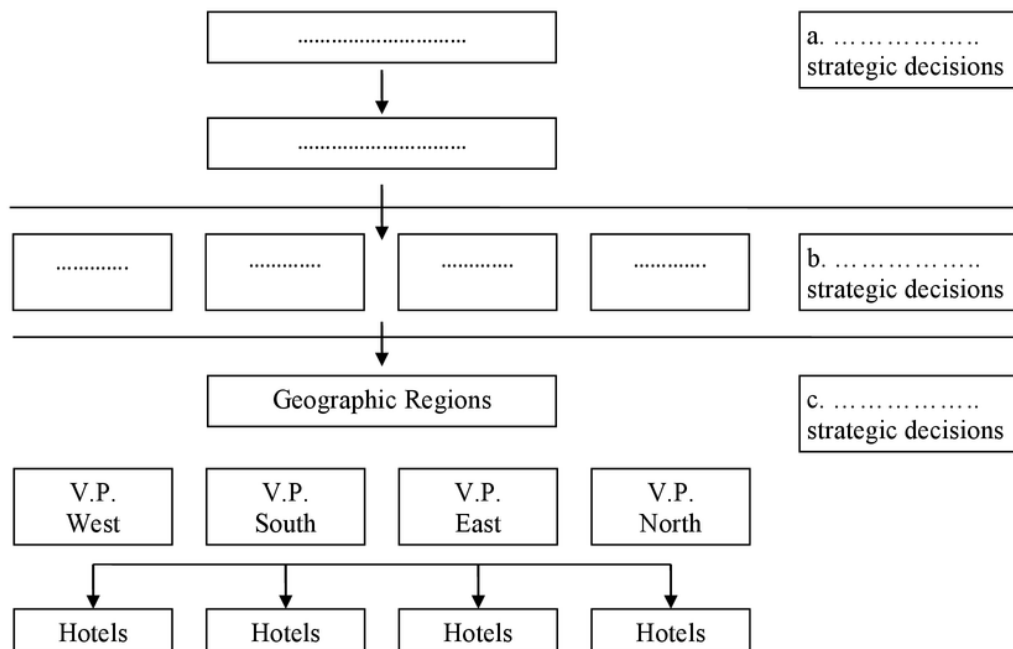
At the top of the management pyramid is the Board of Directors, and the chief executive officer (CEO), who is responsible to the Board of Directors for the business-overall performance of the company. At this level within a hotel company, business-level strategic decisions are made. If this business were part of a larger conglomerate, corporate level strategic decisions would be made at an even higher level. If the company is not part of a larger corporation, corporate and business-level decisions are one and the same and are made by the Board of Directors and the CEO.

Directly below the CEO are the functional executive of the hotel company, so named because each is responsible for a particular specialized aspect, or function, of the business such as marketing, accounting, personnel, and food and beverage. Included at this level is a vice-president of operations who has overall responsibility for the operational activities of the company's hotels. Functional level strategic decisions are made at this level in the management structure.

In a relatively small hotel company, the individual hotel managers report directly to the vice-president of hotel operations. In larger companies, regional vice presidents oversee the operations of individual hotels in various geographic regions. They in turn report to the V.P. of hotel operations. At the regional and individual hotel level, operational strategic decisions are made

14. Based on the reading text above, supply the blank spaces in the flowchart below!





### Text 5

*Nabe*, needless to say, is just one of the countless Japanese foods that have not made journey overseas. Another very popular everyday food that falls into this category is Japanese noodles. As with *nabe*, exactly how you eat your noodles often depend upon where in Japan you eat them. Generally speaking, people in the east favor *soba* noodles, which are made from buckwheat flour kneaded into five brown threads, while those in the west prefer thick wheat *udon* noodles. Both *soba* and *udon* vary in the taste and thickness according to the region in which they are made. Most commonly served in a very simple, hot soy-sauce based broth or plain-boiled to be dipped into liquid condiment, Japanese noodles are a healthy nourishing treat. Here again, though, the foreign influence in Japanese culinary habits is very much in evidence, with curry *soba* and curry *udon* being among the most popular Japanese noodles variants.

15. In Japan, what class of food does *nabe* belong to?
16. What are actually *soba* and *udon*?
17. Why are *soba* and *udon* different?
18. How are *soba* and *udon* served?
19. What foreign influence is there existing in both *soba* and *udon*?
20. If we go to cities in eastern regions of Japan, say Sendai, Yokohama, Chiba or Suzuka, which kind of Nabe do you think we will find?
21. Unlike *shabu-shabu* or *teriyaki*, why do we, foreigners, not recognize *nabe* as Japanese food?

### Text 6

During the General Manager (GM) research, I spent time with executives and staff of the outstanding, excellently managed hotels. It was rather remarkable to discover that the management and staff of these hotels felt that their hotel was by far the best in the city! The reasons they gave for being the best, to be sure, varied from one hotel to another. Still, this feeling of being the best was widespread among the management and staff and unmistakable in each hotel. The GMs of these hotels had developed among their management and staff a self-concept, or culture that said a lot about how that organization felt the hotel should be run. Call it an operating philosophy a self-concept, or a culture; whatever it is called, it's quite clear that developing the right kind of culture is an important goal for all well managed hotels.

22. What profession do you think is the writer of the text?
23. What for did the writer spend time with management of good hotels?
24. What do you think is lacking in the hotel with poorly managed hotels?
25. Why is the claim for being 'the best hotel' made by hotel management important?

### Text 7

Eight of ten general managers (GMs) that were studied feel that understanding people is one of their main strengths. Furthermore, they considered their ability to develop people as the major reasons for their success. The fifty-three hotel department heads who were interviewed ranked employees' motivation as their greatest challenge. Some seem to be natural when it comes to people. William Scully feels that understanding and being able to deal with all kinds of people is one of his main strengths. It turns out that his subordinates feel the same way! They refer to him as fair, approachable, able to handle all kinds of people, caring, a friend, a strong leader with an easy manner, and a person who makes you want to do a good job for him. Frank Anderson is an extrovert with a different saying for every occasion. Henni Le Sassier feels that his elegance and cleverness with people springs from an artistic mother.

26. Mention sentences in the text above that have words or phrases that reflect facts!
27. Which sentences in the text above that have words or phrases that reflect opinions?
28. According to the text above, there are factors that make successful leaders. Mention the factors according to general managers (GMs)!
29. According to Frank Anderson, a department head who is a quiet person will be successful in his profession. Give your comment!

### Text 8

The success of *Karihana's* book has no doubt been welcomed by the Japanese government, which he also been working to promote Japanese food overseas. Since being set up in 2004, the government's Intellectual Property Strategy Headquarters has followed a policy trying to raise the brand value and push global sales of Japanese food. Despite the good reputation of Japanese food, there are still comparatively few Japanese restaurants

outside Japan. Only by increasing opportunities for people to sample genuine Japanese cooking will Japanese food ever fully establish itself as a brand in the way that Italian or Mexican or Indian food has.

30. Supply the best title for the text above!
31. Who is possibly *Karihara*?
32. Why is the Japanese government pleased with the success of *Karihara's* book?
33. What kind of success has the book made?
34. With the success, will the Japanese government depend fully on *Karihara's* book to introduce Japanese food overseas?
35. Has the book been written on behalf of the government?
36. What advantages will probably *Karihara* get from the success?
37. What effect does the book have on the sale of Japanese food in Japan?
38. According to the writer, what requirements need to be fulfilled to make Japanese food recognized as a brand abroad?
39. Why has Italian or Mexican or Indian food been well known everywhere?

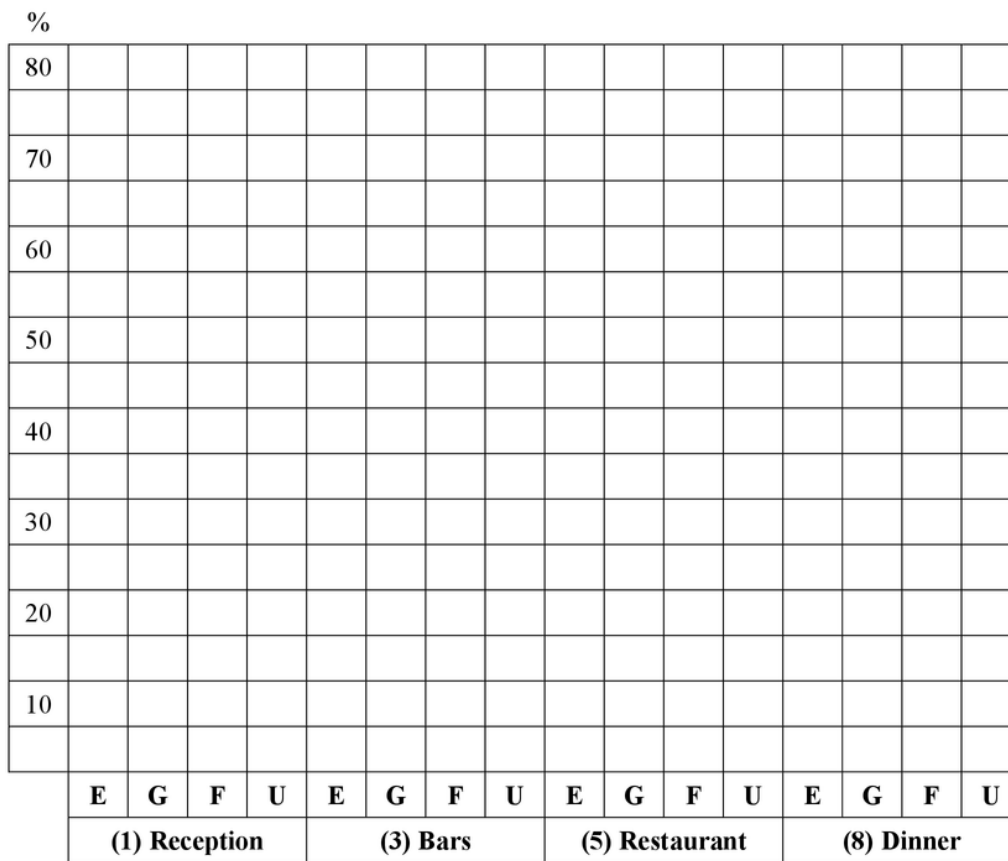
#### Text 9

The following is the summary of the analysis to the answers to a feedback questionnaire in a survey that ***Flamboyant International Hotel***, Batu has recently had to 1521 hotel guests during their stay in the hotel in 2005

#### Summary of the Analysis to the Answers to a Feedback Questionnaire

No	Questions for Area of Services	Responses (%)			
		Excellent	Good	Fair	Unacceptable
1	Reception	0	15	75	15
2	Lobby	0	20	50	30
3	Bars	15	40	45	0
4	Restaurant - quality of environment	50	30	20	0
5	Restaurant - quality of service	10	15	70	5
6	Breakfast - quality of food	50	30	15	5
7	Lunch - quality of food	70	10	20	0
8	Dinner - quality of food	60	30	10	0
9	Bedrooms	15	60	20	5
		<b>Yes</b>		<b>No</b>	
10	Having Experienced the hotel facilities of your stay, would you choose to stay in this hotel again in the future?	35		65	
11	As a result of service and hospitality that you have received during your stay, would you choose to stay in this hotel again in the future?	58		42	

40. Plot the results for the following areas: (1) **Reception**, (3) **Bars**, (5) **Restaurant - quality of service**, and (8) **Dinner - quality of food** in the form of bar charts using the following visual presentation:



**Note:** E = Excellent; G = Good; F = Fair; U = Unacceptable

41. With the result as shown in the table above, do you think that the management of ***Flamboyant International Hotel*** will be satisfied?
42. In which areas do you think were the results most disappointing?

### Appendix 3: Moderators' Evaluation Feedback Sheet

Please provide your evaluative feedback on the test focusing on 8 (eight) aspects that follow:

- |                       |                         |                 |               |
|-----------------------|-------------------------|-----------------|---------------|
| 1. Agreement          | A. Unrelated            | B. Fair         | C. Close      |
|                       | D. Other remarks: _____ |                 |               |
|                       | _____                   |                 |               |
| 2. Item Wording       | A. Clear                | B. Fairly clear | C. Very clear |
|                       | D. Other remarks: _____ |                 |               |
|                       | _____                   |                 |               |
| 3. Test function      | A. Bad                  | B. Fair         | C. Good       |
|                       | D. Other remarks: _____ |                 |               |
|                       | _____                   |                 |               |
| 4. Presentation Order | A. Bad                  | B. Fair         | C. Good       |
|                       | D. Other remarks: _____ |                 |               |
|                       | _____                   |                 |               |
| 5. Instructions       | A. Unclear              | B. Fairly clear | C. Very clear |
|                       | D. Other remarks: _____ |                 |               |
|                       | _____                   |                 |               |
| 6. Difficulty         | A. Easy                 | B. Moderate     | C. Difficult  |
|                       | D. Other remarks: _____ |                 |               |
|                       | _____                   |                 |               |
| 7. Time Allocation    | A. Little               | B. Fair         | C. Much       |
|                       | D. Other remarks: _____ |                 |               |
|                       | _____                   |                 |               |
|                       | _____                   |                 |               |

8. Test Layout

A. Bad

B. Fair

C. Good

D. Other remarks: \_\_\_\_\_

\_\_\_\_\_

**Other comments : (Please specify here)**

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

**Moderator:** \_\_\_\_\_ **Signature:** \_\_\_\_\_

**Teaching Experience:** \_\_\_\_\_ **Date:** \_\_\_\_\_

#### Appendix 4: Test Paper for Informal Trial/Pilot Testing

##### **Directions:**

1. There are 8 short texts in this test followed with 35 questions in all;
2. Do all the test items following these texts by answering all the questions after each text, or by filling in the blank spaces provided in the task;
3. Write your answer in the answer sheet provided;
4. Answer as briefly as possible, max. **one short** sentence;
5. Work by yourself;
6. Do not open your dictionary.

##### **Text 1**

Hotels, both large and small, are exciting places. Many of the people who work in hotels are drawn to them because of this excitement. It is exciting, for example, to attend a gala banquet in a fine hotel. I remember attending a testimonial dinner one evening. 400 guests, all in evening attire, were having cocktails in an area adjacent to the banquet room. When it was time to be seated, the partitioned wall separating the guests from the banquet room was retracted revealing the splendidly arranged tables. The sight was so beautiful that the 400 guests spontaneously began to applaud.

1. What is the writer's opinion about hotels in the text above?
2. What is the purpose of the writer in presenting his/her personal experience in attending a gala banquet?
3. What is the best title for the text above?

##### **Text 2**

Groups of people in hospitality industries, both formal and informal, and cliques (subgroups) are naturally formed within most hospitality operations. When they are effectively managed, they can be built into teams that will yield very effective group performance. Managers must know how to work with groups. This is important when supervising employees on the job and when providing training to staff members. Group training, during which the trainer serves as a facilitator (helper) of the group members' learning activities, involves application of the same basic principles as does working with groups in on-the-job performance of work tasks. Managers, then, must know how to effectively develop and work with groups when supervising and when training.

4. According to the writer, what are the responsibilities of managers?
5. What does the writer mean by basic principles (line 3 from bottom) in the text?
6. A manager needs to work together with staff members on the job. Give an example!

##### **Text 4**

It is perhaps easiest to grasp the hierarchy of strategic decision making by taking a look at a typical organization chart for a hotel company as shown by a figure below. The



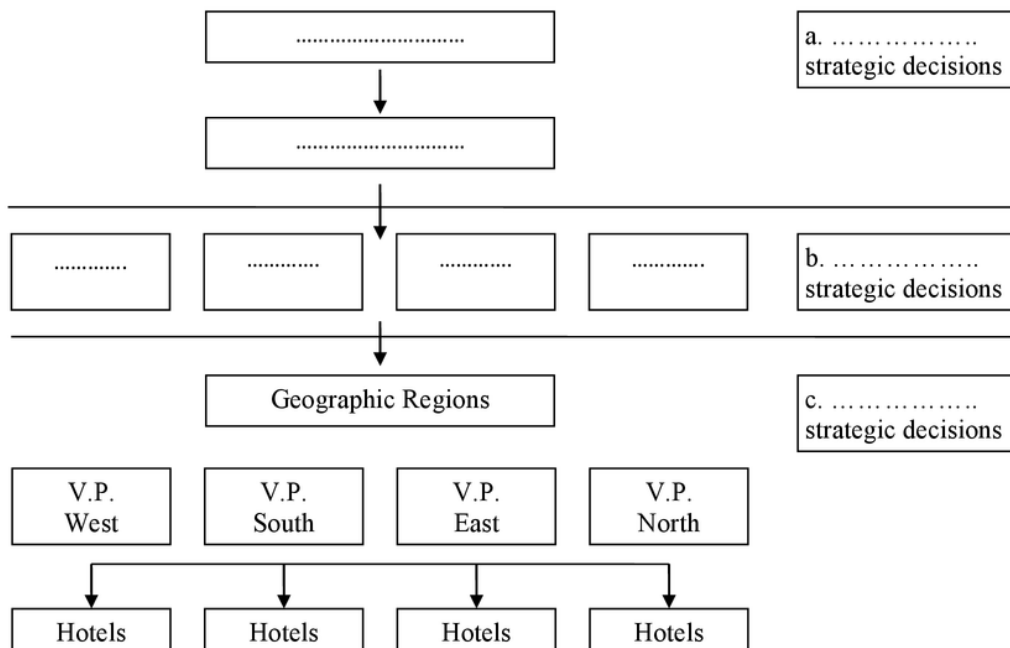
figure shows such an organization chart along with the kinds of strategic decisions that are normally made at various levels within the organization.

At the top of the management pyramid is the Board of Directors, and the chief executive officer (CEO), who is responsible to the Board of Directors for the business-overall performance of the company. At this level within a hotel company, business-level strategic decisions are made. If this business were part of a larger conglomerate, corporate level strategic decisions would be made at an even higher level. If the company is not part of a larger corporation, corporate and business-level decisions are one and the same and are made by the Board of Directors and the CEO.

Directly below the CEO are the functional executive of the hotel company, so named because each is responsible for a particular specialized aspect, or function, of the business such as marketing, accounting, personnel, and food and beverage. Included at this level is a vice-president of operations who has overall responsibility for the operational activities of the company's hotels. Functional level strategic decisions are made at this level in the management structure.

In a relatively small hotel company, the individual hotel managers report directly to the vice-president of hotel operations. In larger companies, regional vice presidents oversee the operations of individual hotels in various geographic regions. They in turn report to the V.P. of hotel operations. At the regional and individual hotel level, operational strategic decisions are made.

7. Based on the reading text above, supply the blank spaces in the flowchart below!



### Text 5

*Nabe*, needless to say, is just one of the countless Japanese foods that have not made journey overseas. Another very popular everyday food that falls into this category is Japanese noodles. As with *nabe*, exactly how you eat your noodles often depend upon where in Japan you eat them. Generally speaking, people in the east favor *soba* noodles, which are made from buckwheat flour kneaded into five brown threads, while those in the west prefer thick wheat *udon* noodles. Both *soba* and *udon* vary in the taste and thickness according to the region in which they are made. Most commonly served in a very simple, hot soy-sauce based broth or plain-boiled to be dipped into liquid condiment, Japanese noodles are a healthy nourishing treat. Here again, though, the foreign influence in Japanese culinary habits is very much in evidence, with curry *soba* and curry *udon* being among the most popular Japanese noodles variants.

8. In Japan, what class of food does *nabe* belong to?
9. What are actually *soba* and *udon*?
10. Why are *soba* and *udon* different?
11. How are *soba* and *udon* served?
12. What foreign variants affect both *soba* and *udon*?
13. If we go to cities in eastern regions of Japan, say Sendai, Yokohama, Chiba or Suzuka, which kind of Nabe do you think we will find?
14. Why do we, foreigners, not recognize *nabe* as Japanese food like *shabu-shabu* or *teriyaki*?

### Text 6

During the General Manager (GM) research, I spent time with executives and staff of the outstanding, excellently managed hotels. It was rather remarkable to discover that the management and staff of these hotels felt that their hotel was by far the best in the city! The reasons they gave for being the best, to be sure, varied from one hotel to another. Still, this feeling of being the best was widespread among the management and staff and unmistakable in each hotel. The GMs of these hotels had developed among their management and staff a self-concept, or culture that said a lot about how that organization felt the hotel should be run. Call it an operating philosophy, a self-concept, or a culture; whatever it is called, it's quite clear that developing the right kind of culture is an important goal for all well managed hotels.

15. What is the profession of the writer of the text?
16. What is the purpose of the writer by spending time with management of good hotels?
17. What is lacking in the hotel with poorly managed hotels?
18. Why is the claim for being 'the best hotel' made by hotel management important?

### Text 7

Eight of ten general managers (GMs) that were studied feel that understanding people is one of their main strengths. Furthermore, they considered their ability to develop people as

the major reasons for their success. The fifty-three hotel department heads who were interviewed ranked employees' motivation as their greatest challenge. Some seem to be natural when it comes to people. William Scully feels that understanding and being able to deal with all kinds of people is one of his main strengths. It turns out that his subordinates feel the same way! They refer to him as fair, approachable, able to handle all kinds of people, caring, a friend, a strong leader with an easy manner, and a person who makes you want to do a good job for him. Frank Anderson is an extrovert with a different saying for every occasion. Henri Le Sassier feels that his elegance and cleverness with people springs from an artistic mother.

19. Which sentences in the text above contain facts?
20. Which sentences in the text above reflect the writer's opinions?
21. According to general managers (GMs) in the text above, what make successful leaders?
22. Do you agree with Frank Anderson that a department head who is a quiet person will be successful in his profession? Give your comment!

#### Text 8

The success of *Karihard's* book has no doubt been welcomed by the Japanese government, which he also been working to promote Japanese food overseas. Since being set up in 2004, the government's Intellectual Property Strategy Headquarters has followed a policy trying to raise the brand value and push global sales of Japanese food. Despite the good reputation of Japanese food, there are still comparatively few Japanese restaurants outside Japan. Only by increasing opportunities for people to sample genuine Japanese cooking will Japanese food ever fully establish itself as a brand in the way that Italian or Mexican or Indian food has.

23. Supply the best title for the text above!
24. Who is possibly *Karihard*?
25. Why is the Japanese government pleased with the success of *Karihard's* book?
26. What kind of success has the book made?
27. With the success, will the Japanese government depend fully on *Karihard's* book to introduce Japanese food overseas?
28. Has the book been written on behalf of the government?
29. What advantages will probably *Karihard* get from the success?
30. What effect does the book have on the sale of Japanese food in Japan?
31. According to the writer, what requirements need to be fulfilled to make Japanese food recognized as a brand abroad?
32. Why has Italian or Mexican or Indian food been well known everywhere?

#### Text 9

The following is the summary of the analysis to the answers to a feedback questionnaire in a survey that ***Flamboyant International Hotel***, Batu has recently had to 1521 hotel guests during their stay in the hotel in 2005.

**Summary of the Analysis to the Answers to a Feedback Questionnaire**

No	Questions for Area of Services	Responses (%)			
		Excellent	Good	Fair	Unacceptable
1	Reception	0	15	75	15
2	Lobby	0	20	50	30
3	Bars	15	40	45	0
4	Restaurant - quality of environment	50	30	20	0
5	Restaurant - quality of service	10	15	70	5
6	Breakfast - quality of food	50	30	15	5
7	Lunch - quality of food	70	10	20	0
8	Dinner - quality of food	60	30	10	0
9	Bedrooms	15	60	20	5
		<b>Yes</b>		<b>No</b>	
10	Having experienced the hotel facilities of your stay, would you choose to stay in this hotel again in the future?	35		65	
11	As a result of service and hospitality that you have received during your stay, would you choose to stay in this hotel again in the future?	58		42	

33. Plot the for the following areas: (1) **Reception**, (3) **Bars**, (5) **Restaurant – quality of service**, and (8) **Dinner – quality of food** in the form of bar charts using the following visual presentation:

%																	
80																	
70																	
60																	
50																	
40																	
30																	
20																	
10																	
	<b>E</b>	<b>G</b>	<b>F</b>	<b>U</b>	<b>E</b>	<b>G</b>	<b>F</b>	<b>U</b>	<b>E</b>	<b>G</b>	<b>F</b>	<b>U</b>	<b>E</b>	<b>G</b>	<b>F</b>	<b>U</b>	
	<b>(1) Reception</b>				<b>(3) Bars</b>				<b>(5) Restaurant</b>				<b>(8) Dinner</b>				

**Note:** E = Excellent; G = Good; F = Fair; U = Unacceptable

34. With the result as shown in the table above, do you think that the management of ***Flamboyant International Hotel*** will be satisfied?
35. In which areas do you think were the results most disappointing?

#### Appendix 5: Students' Evaluation Feedback Sheet

Berikan pendapat Anda tentang tes **Bahasa Inggris IV** yang baru saja Anda alami ditinjau dari 7 (tujuh) hal berikut:

- |                     |                         |                |                 |
|---------------------|-------------------------|----------------|-----------------|
| 1. Instruksi        | A. Tidak jelas          | B. Cukup jelas | C. Jelas sekali |
|                     | D. Other remarks: _____ |                |                 |
| _____               |                         |                |                 |
| 2. Pertanyaan       | A. Tidak jelas          | B. Cukup jelas | C. Jelas sekali |
|                     | D. Other remarks: _____ |                |                 |
| _____               |                         |                |                 |
| 3. Kesukaran        | A. Mudah                | B. Sedang      | C. Sulit        |
|                     | D. Other remarks: _____ |                |                 |
| _____               |                         |                |                 |
| 4. Alokasi waktu    | A. Kurang               | B. Cukup       | C. Berlebih     |
|                     | D. Other remarks: _____ |                |                 |
| _____               |                         |                |                 |
| 5. Tata Letak       | A. Jelek                | B. Cukup       | C. Bagus        |
|                     | D. Other remarks: _____ |                |                 |
| _____               |                         |                |                 |
| 6. Urutan Kesukaran | A. Jelek                | B. Cukup       | C. Bagus        |
|                     | D. Other remarks: _____ |                |                 |
| _____               |                         |                |                 |
| 7. Pengaruh         | A. Tertekan             | B. Biasa       | C. Tertantang   |
|                     | D. Other remarks: _____ |                |                 |
| _____               |                         |                |                 |

**Terimakasih.**

Appendix 6: Respondents' Responses to All the Test Items during Informal Trial

Item No	Respondents						Total
	1	2	3	4	5	6	
1	1	1	1	1	1	0	5
2	0	0	0	0	0	0	0
3	1	0	1	0	0	1	3
4	1	1	1	1	1	0	5
5	1	0	1	0	1	0	3
6	0	0	1	0	0	0	1
7	0	0	0	1	1	0	2
8	0	0	1	1	1	0	3
9	1	1	1	1	1	0	5
10	0	1	1	0	0	0	2
11	1	1	1	1	1	0	5
12	0	1	0	0	1	0	2
13	0	0	1	1	1	0	3
14	0	0	0	0	1	0	1
15	0	0	0	0	0	0	0
16	0	0	1	0	0	0	1
17	0	0	0	0	0	0	0
18	0	0	0	0	1	0	1
19	0	1	0	0	0	0	1
20	0	0	0	0	0	0	0
21	1	1	0	0	0	1	3
22	0	1	0	0	0	0	1
23	0	0	0	0	0	0	0
24	0	1	0	0	0	0	1
25	0	0	0	0	1	0	1
26	1	0	0	0	0	0	1
27	0	0	0	0	1	0	1
28	1	0	0	0	1	0	2
29	0	0	0	0	0	0	0
30	0	0	0	0	0	0	0
31	1	0	0	0	1	0	2
32	0	0	0	0	0	0	0
33	0	0	0	0	0	0	0
34	0	0	0	0	0	0	0
35	1	0	0	0	0	0	1
Total	11	10	11	7	15	2	

Note: 1 = correct answer; 0 = incorrect answer



**UJIAN TENGAH SEMESTER GANJIL 2005/2006  
PROGRAM DIPLOMA III PARIWISATA  
UNIVERSITAS MERDEKA MALANG**

<b>Mata Ujian</b>	<b>: Bhs. Inggris IV</b>	<b>Smt./Kelas</b>	<b>: V/A, B, C</b>
<b>Dosen</b>	<b>: Tim</b>	<b>Jurusan</b>	<b>: PHT</b>
<b>Hari</b>	<b>: Rabu</b>	<b>Waktu</b>	<b>: 90 menit</b>
<b>Tanggal</b>	<b>: 23 Nopember 2005</b>	<b>Sifat</b>	<b>: Tutup Buku</b>

---

**Directions:**

- There are 8 short texts in this test followed with 35 questions in all;
  - Do all the test items following these texts by answering all the questions after each text, or by filling in the blank spaces provided in the task;
  - Write your answer in the answer sheet provided;
  - Answer as briefly as possible, max. **one short** sentence;
  - Work by yourself
  - Do not open your dictionary.
- 

**Text 1**

Hotels, both large and small, are exciting places. Many of the people who work in hotels are drawn to them because of this excitement. It is exciting, for example, to attend a gala banquet in a fine hotel. I remember attending a testimonial dinner one evening. 400 guests, all in evening attire, were having cocktails in an area adjacent to the banquet room. When it was time to be seated, the partitioned wall separating the guests from the banquet room was retracted revealing the splendidly arranged tables. The sight was so beautiful that the 400 guests spontaneously began to applaud.

1. What is the writer's opinion about hotels in the text above?
2. What is the best title for the text above?

**Text 2**

Groups of people in hospitality industries, both formal and informal, and cliques (subgroups) are naturally formed within most hospitality operations. When they are effectively managed, they can be built into teams that will yield very effective group performance. Managers must know how to work with groups. This is important when supervising employees on the job and when providing training to staff members. Group training, during which the trainer serves as a facilitator (helper) of the group members' learning activities, involves application of the same basic principles as does working with groups in on-the-job performance of work tasks. Managers, then, must know how to effectively develop and work with groups when supervising and when training.

3. According to the writer, what are the responsibilities of managers?

4. What does the writer mean by basic principles (line 3 from bottom) in the text?
5. A manager needs to work together with staff members on the job. Give an example!

### Text 3

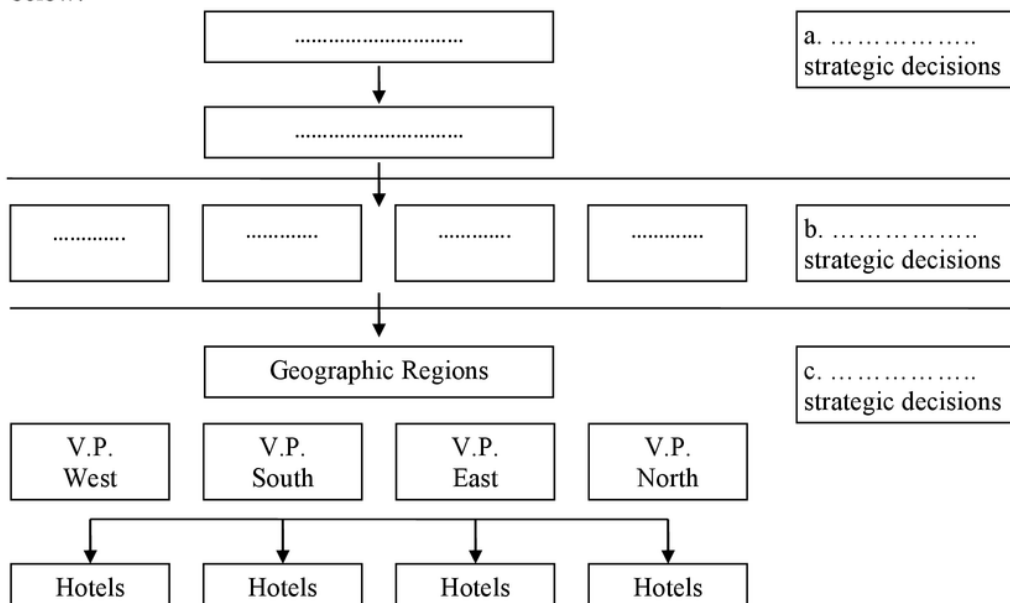
It is perhaps easiest to grasp the hierarchy of strategic decision making by taking a look at a typical organization chart for a hotel company as shown by a figure below. The figure shows such an organization chart along with the kinds of strategic decisions that are normally made at various levels within the organization.

At the top of the management pyramid is the Board of Directors, and the chief executive officer (CEO), who is responsible to the Board of Directors for the business-overall performance of the company. At this level within a hotel company, business-level strategic decisions are made. If this business were part of a larger conglomerate, corporate level strategic decisions would be made at an even higher level. If the company is not part of a larger corporation, corporate and business-level decision are one and the same and are made by the Board of Directors and the CEO.

Directly below the CEO are the functional executive of the hotel company, so named because each is responsible for a particular specialized aspect, or function, of the business such as marketing, accounting, personnel, and food and beverage. Included at this level is a vice-president of operations who has responsibility for the operational activities of the company's hotels. Functional level strategic decisions are made at this level in the management structure.

In a relatively small hotel company, the individual hotel managers report directly to the vice-president of hotel operations. In larger companies, regional vice presidents oversee the operations of individual hotels in various geographic regions. They in turn report to the V.P. of hotel operations. At the regional and individual hotel level, operational strategic decisions are made.

6. Based on the reading text above, supply the blank spaces in the flowchart below!



#### Text 4

*Nabe*, needless to say, is just one of the countless Japanese foods that have not made journey overseas. Another very popular everyday food that falls into this category is Japanese noodles. As with *nabe*, exactly how you eat your noodles often depend upon where in Japan you eat them. Generally speaking, people in the east favor *soba* noodles, which are made from buckwheat flour kneaded into five brown threads, while those in the west prefer thick wheat *udon* noodles. Both *soba* and *udon* vary in the taste and thickness according to the region in which they are made. Most commonly served in a very simple, hot soy-sauce based broth or plain-boiled to be dipped into liquid condiment, Japanese noodles are a healthy nourishing treat. Here again, though, the foreign influence in Japanese culinary habits is very much in evidence, with curry *soba* and curry *udon* being among the most popular Japanese noodles variants.

7. In Japan, what class of food does *nabe* belong to?
8. What are actually *soba* and *udon*?
9. Why are *soba* and *udon* different?
10. How are *soba* and *udon* served?
11. What foreign variants affect both *soba* and *udon*?
12. If we go to cities in eastern regions of Japan, say Sendai, Yokohama, Chiba or Suzuka, which kind of Nabe do you think we will find?
13. Why do we, foreigners, not recognize *nabe* as Japanese food like *shabu-shabu* or *teriyaki*?

#### Text 5

During the General Manager (GM) research, I spent time with executives and staff of the outstanding, excellently managed hotels. It was rather remarkable to discover that the management and staff of these hotels felt that their hotel was by far the best in the city! The reasons they gave for being the best, to be sure, varied from one hotel to another. Still, this feeling of being the best was widespread among the management and staff and unmistakable in each hotel. The GMs of these hotels had developed among their management and staff a self-concept, or culture that said a lot about how that organization felt the hotel should be run. Call it an operating philosophy, a self-concept, or a culture; whatever it is called, it's quite clear that developing the right kind of culture is an important goal for all well managed hotels.

14. What is the purpose of the writer by spending time with management of good hotels?
15. What is lacking in the hotel with poorly managed hotels?
16. Why is the claim for being 'the best hotel' made by hotel management important?

#### Text 6

Eight of ten general managers (GMs) that were studied feel that understanding people is one of their main strengths. Furthermore, they considered their ability to develop people as the major reasons for their success. The fifty-three hotel department heads who were

interviewed ranked employees' motivation as their greatest challenge. Some seem to be natural when it comes to people. William Scully feels that understanding and being able to deal with all kinds of people is one of his main strengths. It turns out that his subordinates feel the same way! They refer to him as fair, approachable, able to handle all kinds of people, caring, a friend, a strong leader with an easy manner, and a person who makes you want to do a good job for him. Frank Anderson is an extrovert with a different saying for every occasion. Henri Le Sassier feels that his elegance and cleverness with people springs from an artistic mother.

17. Which sentences in the text above that contain facts?
18. Which sentences in the text above that reflect the writer's opinions?
19. According to general managers (GMs) in the text above, what make successful leaders?
20. Do you agree with Frank Anderson that a department head who is a quiet person will be successful in his profession? Give your comment!

#### Text 7

The success of *Karihara's* book has no doubt been welcomed by the Japanese government, which he also been working to promote Japanese food overseas. Since being set up in 2004, the government's Intellectual Property Strategy Headquarters has followed a policy trying to raise the brand value and push global sales of Japanese food. Despite the good reputation of Japanese food, there are still comparatively few Japanese restaurants outside Japan. Only by increasing opportunities for people to sample genuine Japanese cooking will Japanese food ever fully establish itself as a brand in the way that Italian or Mexican or Indian food has.

21. Who is possibly *Karihara*?
22. Why is the Japanese government pleased with the success of *Karihara's* book?
23. What kind of success has the book made?
24. With the success, will the Japanese government depend fully on *Karihara's* book to introduce Japanese food overseas?
25. Has the book been written on behalf of the government?
26. What advantages will probably *Karihara* get from the success?
27. According to the writer, what requirements need to be fulfilled to make Japanese food recognized as a brand abroad?
28. Why has Italian or Mexican or Indian food been well known everywhere?

#### Text 8

The following is the summary of the analysis to the answers to a feedback questionnaire in a survey that ***Flamboyant International Hotel***, Batu has recently had to 1521 hotel guests during their stay in the hotel in 2005.

**Summary of the Analysis to the Answers to a Feedback Questionnaire**

No	Questions for Area of Services	Responses (%)			
		Excellent	Good	Fair	Unacceptable
1	Reception	0	15	75	15
2	Lobby	0	20	50	30
3	Bars	15	40	45	0
4	Restaurant - quality of environment	50	30	20	0
5	Restaurant - quality of service	10	15	70	5
6	Breakfast - quality of food	50	30	15	5
7	Lunch - quality of food	70	10	20	0
8	Dinner - quality of food	60	30	10	0
9	Bedrooms	15	60	20	5
		<b>Yes</b>		<b>No</b>	
10	Having experienced the hotel facilities of your stay, would you choose to stay in this hotel again in the future?	35		65	
11	As a result of service and hospitality that you have received during your stay, would you choose to stay in this hotel again in the future?	58		42	

29. Plot the results for the following areas: (1) **Reception**, (3) **Bars**, (5) **Restaurant - quality of service**, and (8) **Dinner - quality of food**. Darken the areas of the following visual presentation.

%																	
80																	
70																	
60																	
50																	
40																	
30																	
20																	
10																	
	<b>E</b>	<b>G</b>	<b>F</b>	<b>U</b>	<b>E</b>	<b>G</b>	<b>F</b>	<b>U</b>	<b>E</b>	<b>G</b>	<b>F</b>	<b>U</b>	<b>E</b>	<b>G</b>	<b>F</b>	<b>U</b>	
	<b>(1) Reception</b>				<b>(3) Bars</b>				<b>(5) Restaurant</b>				<b>(8) Dinner</b>				

**Note:** E = Excellent; G = Good; F = Fair; U = Unacceptable

30. With the result as shown in the table above, do you think that the management of ***Flamboyant International Hotel*** will be satisfied?

31. In which areas do you think were the results most disappointing?

Appendix 8: Answer Sheet of the Final Test Form

**UJIAN TENGAH SEMESTER GANJIL 2005/2006  
PROGRAM DIPLOMA III PARIWISATA  
UNIVERSITAS MERDEKA MALANG**

Name : \_\_\_\_\_  
Nomor Ujian/Nirm : \_\_\_\_\_  
Mata kuliah : \_\_\_\_\_  
Tanggal : \_\_\_\_\_  
Tanda Tangan : \_\_\_\_\_

---

**ANSWER SHEET**

Before you answer, please pay attention carefully to the way how to give a short answer as provided in the example below.

Example

Questions	Short Answers
<i>In the hotel, where was the dinner held?</i>	<i>The banquet room</i>
<i>How many people were there present?</i>	<i>400</i>

Now, do all the following test tasks like the ones shown in the example above so that you can concentrate your reading and save your time.

**Text 1**

1. \_\_\_\_\_
2. \_\_\_\_\_

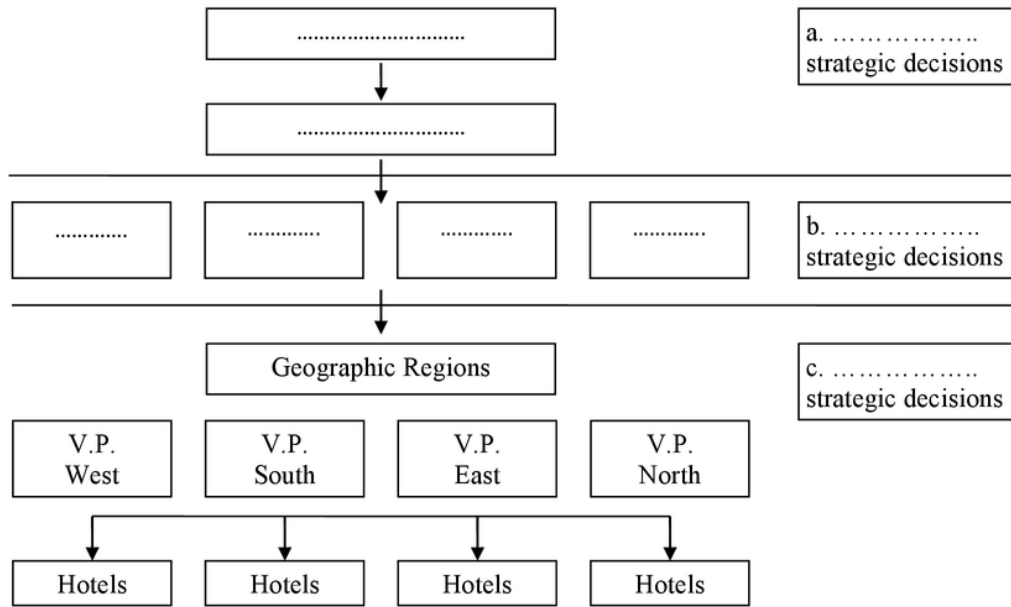
**Text 2**

3. \_\_\_\_\_
4. \_\_\_\_\_
5. \_\_\_\_\_

**Text 3**

6. Complete the diagram that follows based on the information in the text!





#### Text 4

7. \_\_\_\_\_
8. \_\_\_\_\_
9. \_\_\_\_\_
10. \_\_\_\_\_
11. \_\_\_\_\_
12. \_\_\_\_\_
13. \_\_\_\_\_

#### Text 5

14. \_\_\_\_\_
15. \_\_\_\_\_
16. \_\_\_\_\_

#### Text 6

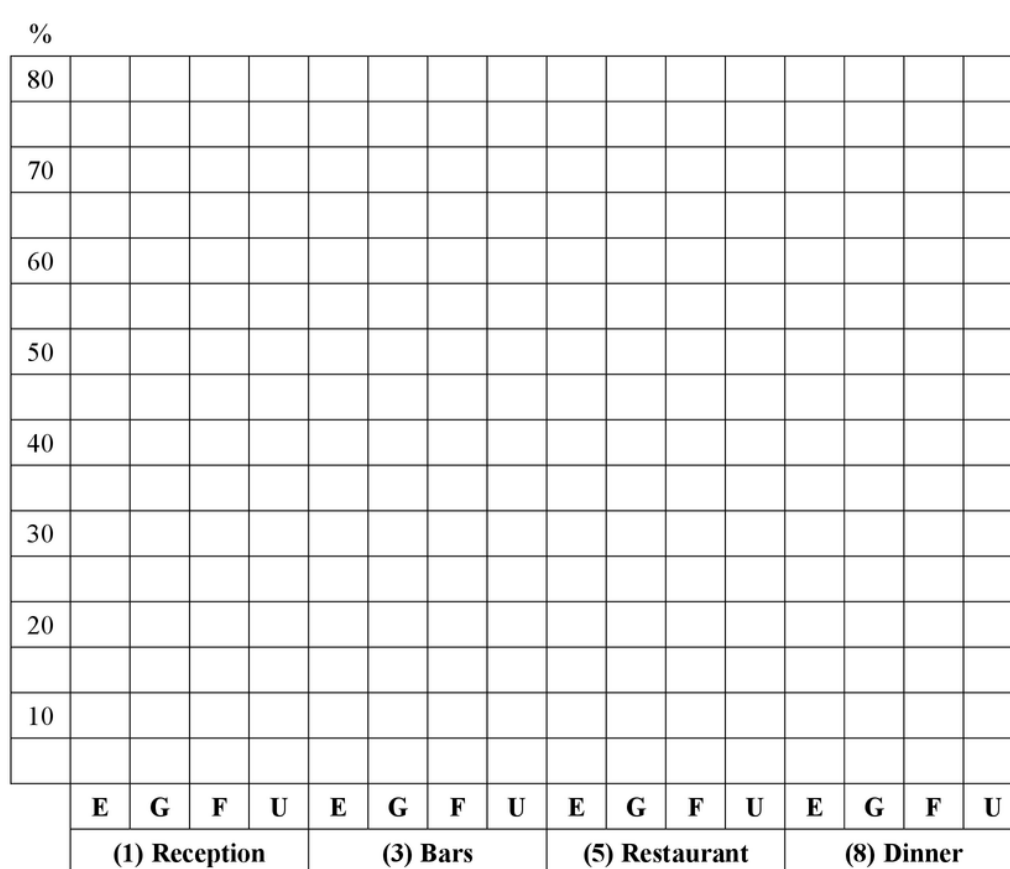
17. \_\_\_\_\_
18. \_\_\_\_\_
19. \_\_\_\_\_
20. \_\_\_\_\_

#### Text 7

21. \_\_\_\_\_
22. \_\_\_\_\_
23. \_\_\_\_\_
24. \_\_\_\_\_
25. \_\_\_\_\_
26. \_\_\_\_\_
27. \_\_\_\_\_
28. \_\_\_\_\_

# Text 9

29. Darken the areas according to the data



**Note:** E = Excellent; G = Good; F = Fair; U = Unacceptable

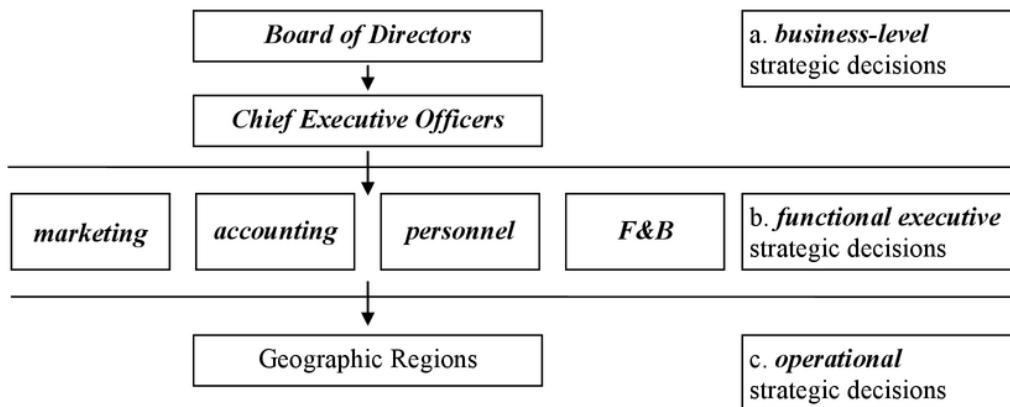
30. \_\_\_\_\_
31. \_\_\_\_\_

# Appendix 9: Final Scoring Guide

No	Score Value	
	0	1
1	No answer or unintelligible written responses	Exciting (places)
2	ditto	Hotel as exciting places
3	ditto	to work with group; and develop group/(/people)
4	ditto	How to work with group
5	ditto	Supervising the work task
6	ditto	See chart attached for No 6 below
7	ditto	Noodles
8	ditto	They are noodles
9	ditto	<i>Soba</i> is the noodles for people in the east and buckwheat flour; <i>Udon</i> in the west thick wheat
10	ditto	Hot-sauce or plain boiled dipped into liquid condiment
11	ditto	curry
12	ditto	<i>Soba</i> (noodles)
13	ditto	It is not yet introduced abroad
14	ditto	To collect data and information about GM
15	ditto	They lack self-concept or culture
16	ditto	It is an operating philosophy
17	ditto	See attachment for No 17 below
18	ditto	See attachment for No 18 below
19	ditto	<ul style="list-style-type: none"> <li>• understanding people</li> <li>• employees' motivation</li> </ul>
20	ditto	No. S/he is not.
21	ditto	The government civil servant/ civil officer
22	ditto	It will help the promotion of Japanese food overseas
23	ditto	Promotion of Japanese food overscas
24	ditto	No it does not
25	ditto	No is not
26	ditto	More than three logical advantages: for example, Fame and financial advantages
27	ditto	Increasing opportunities for foreigners to taste Japanese food
28	ditto	Foreigners have opportunities to taste these kinds of food, ONO
29	ditto	See attachment for No 29 below
30	ditto	No, s/he is not
31	ditto	Restaurant - quality of service

Attachment to Appendix 9

No. 6: all three clusters

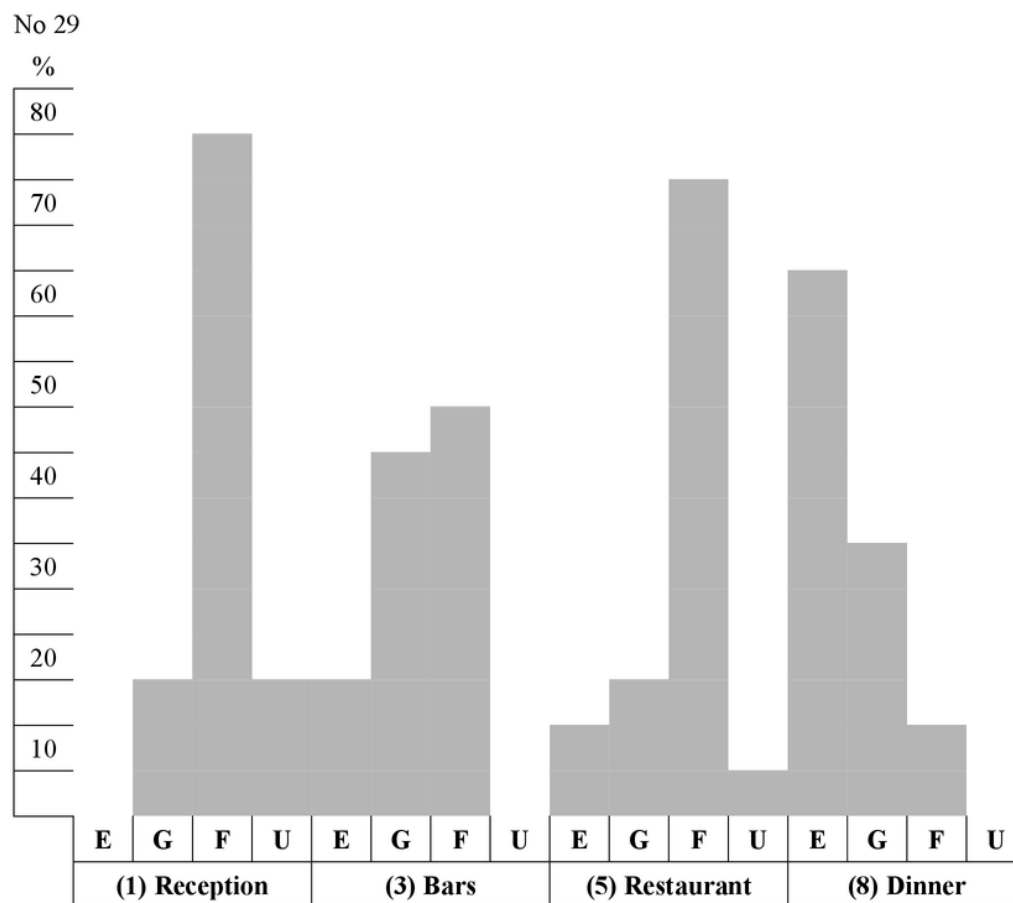


No 17

- Eight of ten general managers (GMs) that were studied feel that understanding people is one of their main strengths.
- The fifty-three hotel department heads who were interviewed ranked employees' motivation as their greatest challenge .

No 18 (Any two)

- They considered their ability to develop people as the major reasons for their success.
- William Scully feels that understanding and being able to deal with all kinds of people is one of his main strengths. They refer to him as fair, approachable, able to handle all kinds of people, caring, a friend, a strong leader with an easy manner, and a person who makes you want to do a good job for him.
- Frank Anderson is an extrovert with a different saying for every occasion
- Henni Le Sassier feels that his elegance and cleverness with people springs from an artistic mother.



# Appendix 10: Raw Score Recording Sheet

Item No	Respondents						Total
	1	2	3	4	5	... etc	
1							
2							
3							
4							
5							
6							
7							
8							
9							
10							
11							
12							
13							
14							
15							
16							
17							
18							
19							
20							
21							
22							
23							
24							
25							
26							
27							
28							
29							
30							
31							
Total							

Note: 1 = correct answer; 0 = incorrect answer

Appendix 11: Students' Data Recording Table

No	Name of Students	Sex	GPA	Total Score	Grade	Other (If any)	Comments
1							
2							
3							
4							
5							
6							
7							
8							
9							
10							
11							
12							
13							
14							
15							
16							
17							
18							
19							
20							
21							
22							
23							
24							
25							
26							
27							
28							
29							
30							
31							
32							
33							
34							
35							
36							
37							
38							
39							
40							



# Appendix 12: Item Analysis Results

Item	U	L	U+L	<i>P</i>	U-L	<i>D</i>
1	25	14	39	.51	11	.28
2	12	1	13	.17	11	.28
3	26	9	35	.46	17	.44
4	13	5	18	.23	8	.21
5	20	5	25	.32	15	.39
6	31	16	37	.61	15	.39
7	25	11	36	.47	14	.36
8	25	8	33	.43	17	.44
9	19	4	23	.3	15	.39
10	16	5	21	.27	11	.28
11.	16	0	16	.21	16	.42
12	13	0	13	.17	13	.34
13	10	0	10	.13	10	.26
14	8	0	8	.13	8	.21
15	8	0	8	.1	8	.21
16	10	0	10	.13	10	.26
17	14	4	18	.23	10	.26
18	8	0	8	.1	8	.21
19	13	0	13	.17	13	.34
20.	7	0	7	.09	7	.18
21	6	2	8	.1	4	.1
22	16	6	22	.28	10	.26
23	4	1	5	.06	3	.07
24	6	2	8	.1	4	.1
25	1	0	1	.01	1	.02
26	0	0	0	0	0	0
27	4	2	6	.07	2	.05
28	7	2	9	.11	5	.13
29	6	0	6	.07	6	.15
30	2	5	7	.09	-3	-.07
31	2	0	2	.03	2	.05

### Appendix 13: Complete Manual Computation

$$\begin{aligned} SS_a &= \frac{1}{N_m} (\sum X_i^2) - \left( \frac{\sum X_i}{N_{ma}} \right)^2 \\ &= \frac{1}{3} (425) - \left( \frac{57}{38} \right)^2 \\ &= 6,291 \end{aligned}$$

$$\begin{aligned} SS_m &= \frac{1}{N_a} (\sum X_j^2) - \left( \frac{\sum X_j}{N_{ma}} \right)^2 \\ &= \frac{1}{8} (1085)^2 - \left( \frac{57}{308} \right)^2 \\ &= 0,25 \end{aligned}$$

$$r = \frac{MSS_a - MSS_c}{MSS_a}$$

$$= \frac{0,899 - 0,237}{0,899}$$

$$= 0,635$$

$$SS_c = 3,083$$

$$\begin{aligned} SS_+ &= (\sum X^2) - \left( \frac{\sum X_i}{N_{ma}} \right)^2 \\ &= 145 - \left( \frac{57}{3 \times 8} \right)^2 \\ &= 9,625 \end{aligned}$$

$$MSS_a = \frac{SS_a}{df_a} \rightarrow \frac{6,291}{7} = 0,8907$$

$$MSS_c = \frac{SS_c}{df_c} \rightarrow \frac{3,083}{13} = 0,237$$

#### Appendix 14: Results of Descriptive Analysis

1. Frequencies: Statistics
2. Frequency Table: Raw Score and Grade

## Frequencies

### Statistics

		RAWScore	GRADE
N	Valid	142	142
	Missing	0	0
Mean		6.0845	15.8076
Std. Deviation		3.2779	10.0987
Variance		10.7446	101.9835
Range		18.00	58.33
Minimum		1.00	1.19
Maximum		19.00	59.32

## Frequency Table

### RAWScore

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1.00	10	7.0	7.0	7.0
	2.00	6	4.2	4.2	11.3
	3.00	16	11.3	11.3	22.5
	4.00	13	9.2	9.2	31.7
	5.00	23	16.2	16.2	47.9
	6.00	13	9.2	9.2	57.0
	7.00	17	12.0	12.0	69.0
	8.00	20	14.1	14.1	83.1
	9.00	8	5.6	5.6	88.7
	10.00	6	4.2	4.2	93.0
	11.00	2	1.4	1.4	94.4
	12.00	2	1.4	1.4	95.8
	13.00	2	1.4	1.4	97.2
	14.00	2	1.4	1.4	98.6
	19.00	2	1.4	1.4	100.0
	Total	142	100.0	100.0	

# GRADE

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1.19	8	5.6	5.6	5.6
	2.38	2	1.4	1.4	7.0
	4.76	6	4.2	4.2	11.3
	5.95	8	5.6	5.6	16.9
	7.14	14	9.9	9.9	26.8
	7.76	2	1.4	1.4	28.2
	8.33	1	.7	.7	28.9
	9.52	10	7.0	7.0	35.9
	10.71	7	4.9	4.9	40.8
	11.90	2	1.4	1.4	42.3
	13.09	6	4.2	4.2	46.5
	14.28	2	1.4	1.4	47.9
	15.47	7	4.9	4.9	52.8
	16.66	2	1.4	1.4	54.2
	17.85	9	6.3	6.3	60.6
	19.04	9	6.3	6.3	66.9
	20.23	6	4.2	4.2	71.1
	20.24	2	1.4	1.4	72.5
	21.42	4	2.8	2.8	75.4
	22.61	6	4.2	4.2	79.6
	22.62	2	1.4	1.4	81.0
	23.80	4	2.8	2.8	83.8
	25.00	3	2.1	2.1	85.9
	26.19	4	2.8	2.8	88.7
	27.38	4	2.8	2.8	91.5
	28.57	2	1.4	1.4	93.0
	30.95	2	1.4	1.4	94.4
	32.14	2	1.4	1.4	95.8
	33.33	2	1.4	1.4	97.2
	36.90	2	1.4	1.4	98.6
	59.52	2	1.4	1.4	100.0
Total		142	100.0	100.0	

## Appendix 15: Result of Factor Analysis

1. KMO and Bartlett's Test
2. Anti-Image Matrices
3. Component Matrix
4. Rotated Component Matrix

### KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.698
Bartlett's Test of Sphericity	Approx. Chi-Square	605.312
	df	120
	Sig.	.000

### Anti-Image Matrices

		ITEM14	ITEM15	ITEM16	ITEM17	ITEM18	ITEM19
Anti-Image Covariance	ITEM1	2.575E-02	-.157	-4.696E-02	2.409E-02	-.230	-4.109E-02
	ITEM2	-.140	5.156E-02	-5.879E-02	3.436E-02	5.455E-02	-7.200E-02
	ITEM6	-.131	-5.356E-02	-9.771E-01	.172	-1.902E-02	-9.694E-02
	ITEM11	6.018E-03	-.193	6.475E-02	.109	-.110	2.082E-02
	ITEM12	-1.383E-03	-.133	-.120	-5.033E-02	4.598E-02	-4.411R-02
	ITEM13	3.304E-02	2.796E-02	1.961E-02	-3.481E-02	-8.585E-02	3.883E-02
	ITEM14	.440	-1.783E-03	-2.541E-03	-.196	-.109	2.646E-02
	ITEM15	-1.783E-03	.592	-7.997E-02	-.119	.144	6.834E-02
	ITEM16	-2.531E-03	-7.997E-02	.639	-2.482E-02	-8.015E-03	1.436E-02
	ITEM17	-.196	-.119	-2.483E-02	.480	-.162	-4.450E-03
	ITEM18	-.109	.144	-8.015E-03	-.162	.571	-4.613E-02
	ITEM19	2.646E-02	6.834E-02	1.436E-02	-4.450E-03	-4.613E-02	.571
	ITEM20	6.059E-02	-4.083E-02	-3.485E-02	-3.649E-02	-2.566E-03	-.227
	ITEM27	-.185	-4.345E-02	-.119	7.762E-02	-1.692E-02	5.348E-02
ITEM28	2.168E-02	-2.619E-02	-.101	-.151	.148	-.197	
ITEM29	-3.037E-02	-6.682E-03	-9.305E-02	5.540E-03	.104	-.246	
Anti-Image Correlation	ITEM1	4.422E-02	-.233	-6.689E-02	3.960E-02	-.347	-6.190E-02
	ITEM2	-.276	8.741E-02	-9.592E-02	6.468E-02	9.414E-02	-.124
	ITEM6	-.231	-8.137E-02	-.143	.289	-2.942E-02	-.150
	ITEM11	1.130E-02	-.312	.101	.195	-.182	3.429E-02
	ITEM12	-2.799E-03	-.232	-.202	-9.758E-02	8.164E-02	-7.829E-02
	ITEM13	6.283E-02	4.585E-02	3.093E-02	-6.337E-02	-.143	6.477E-02
	ITEM14	.716 <sup>a</sup>	-3.496E-03	-4.795E-03	-.428	-.217	5.280E-02
	ITEM15	-3.496E-03	.690 <sup>a</sup>	-.130	-.224	.249	.118
	ITEM16	-4.795E-03	-.130	.851 <sup>a</sup>	-4.482E-02	-1.327E-02	2.377E-02
	ITEM17	-.428	-.224	-4.482E-02	.642 <sup>a</sup>	-.309	-8.499E-03
	ITEM18	-.217	.249	-1.327E-02	-.309	.569 <sup>a</sup>	-8.076E-02
	ITEM19	5.280E-02	.118	2.377E-02	-8.499E-03	-8.076E-02	.603 <sup>a</sup>
	ITEM20	.109	-6.321E-02	-5.192E-02	-6.273E-02	-4.044E-03	-.358
	ITEM27	-.375	-7.588E-02	-.199	.151	-3.007E-02	9.503E-02
ITEM28	3.849E-02	-4.007E-02	-.148	-.256	.230	-.306	
ITEM29	-5.480E-02	-1.040E-02	-.139	9.572E-03	.164	-.390	



### Anti-Image Matrices

	ITEM1	ITEM2	ITEM6	ITEM11	ITEM12	ITEM13
Anti-Image Covariance	ITEM1	.771	-9.125E-02	7.387E-02	.107	9.993E-03
	ITEM2	-9.125E-02	.588	.100	-7.057E-03	-1.157E-02
	ITEM6	7.387E-02	.100	.732	-4.695E-02	-1.382E-02
	ITEM11	.107	-7.057E-03	-.104	.645	-.175
	ITEM12	9.993E-03	-1.157E-02	-4.695E-02	-.175	.555
	ITEM13	3.619E-03	-.195	-1.382E-02	-2.927E-02	-.112
	ITEM14	2.575E-02	-.140	-.131	6.018E-03	-1.383E-03
	ITEM15	-.157	5.156E-02	-5.356E-02	-.193	-.133
	ITEM16	-4.696E-02	-5.879E-02	-9.771E-02	6.475E-02	-.120
	ITEM17	2.409E-02	3.436E-02	.172	.109	-5.038E-02
	ITEM18	-.230	5.455E-02	-1.902E-02	-.110	4.598E-02
	ITEM19	-4.109E-02	-7.200E-02	-9.694E-02	2.082E-02	-4.411E-02
	ITEM20	4.916E-02	-.160	-2.520E-02	-2.909E-02	1.677E-02
	ITEM27	3.745E-02	2.060E-02	9.185E-02	3.436E-02	-6.245E-02
	ITEM28	-2.937E-03	-2.755E-03	8.163E-04	-3.868E-02	9.996E-02
	ITEM29	-5.256E-02	-3.568E-02	-4.173E-02	-7.132E-02	-6.467E-03
Anti-Image Correlation	ITEM1	.582 <sup>a</sup>	-.135	9.829E-02	.152	1.527E-02
	ITEM2	.135	.755 <sup>a</sup>	.153	-1.146E-02	-2.024E-02
	ITEM6	9.829E-02	.153	.570 <sup>a</sup>	-.151	-7.362E-02
	ITEM11	.152	-1.146E-02	-.151	.633 <sup>a</sup>	-.293
	ITEM12	1.527E-02	-2.024E-02	-7.362E-02	-.293	.801 <sup>a</sup>
	ITEM13	5.197E-03	-.321	-2.037E-02	-4.594E-02	-.190
	ITEM14	4.422E-02	-.276	-.231	1.130E-02	-2.799E-03
	ITEM15	-.233	8.741E-02	-8.137E-02	-.312	-.232
	ITEM16	-6.689E-02	-9.591E-02	-.143	.101	-.202
	ITEM17	3.960E-02	6.468E-02	.289	.195	-9.758E-02
	ITEM18	-.347	9.414E-02	-2.942E-02	-.182	8.164E-02
	ITEM19	-6.190E-02	-.124	-.150	3.429E-02	-7.829E-02
	ITEM20	6.665E-02	-.248	-3.507E-02	-4.313E-02	2.679E-02
	ITEM27	5.727E-02	3.609E-02	.144	5.746E-02	-.113
	ITEM28	-3.936E-03	-4.228E-03	1.123E-03	-5.668E-02	.158
	ITEM29	-7.161E-02	-5.568E-02	-5.836E-02	-.106	-1.038E-02

### Anti-Image Matrices

		ITEM20	ITEM27	ITEM28	ITEM29
Anti-Image Covariance	ITEM1	4.916E-02	3.745E-02	-2.937E-03	-5.256E-02
	ITEM2	-.160	2.060E-02	-2.733E-03	-3.568E-02
	ITEM6	-2.520E-02	9.185E-02	8.163E-04	-4.173E-02
	ITEM11	-2.909E-02	3.436E-02	-3.868E-02	-7.132E-02
	ITEM12	1.677E-02	-6.245E-02	9.996E-02	-6.467E-03
	ITEM13	-3.213E-02	-.152	-5.866E-03	6.285E-03
	ITEM14	6.059E-02	-.185	2.168E-02	-3.037E-02
	ITEM15	-4.083E-02	-4.345E-02	-2.619E-02	-6.682E-03
	ITEM16	-3.485E-02	-.119	-.101	-9.305E-02
	ITEM17	-3.649E-02	7.762E-02	-.151	5.540E-03
	ITEM18	-2.566E-03	-1.692E-02	.148	.104
	ITEM19	-.227	5.348E-02	-.197	-.246
	ITEM20	.705	-3.810E-02	-5.108E-03	.121
	ITEM27	-3.810E-02	.554	-8.125E-02	-7.143E-02
	ITEM28	-5.108E-03	-8.125E-02	.722	7.178E-02
	ITEM29	.121	-7.143E-02	7.178E-02	.698
Anti-Image Correlation	ITEM1	6.665E-02	5.727E-02	-3.936E-03	-7.161E-02
	ITEM2	-.248	3.609E-02	-4.228E-03	-5.568E-02
	ITEM6	-3.507E-02	.144	1.123E-03	-5.836E-02
	ITEM11	-4.313E-02	5.746E-02	-5.668E-02	-.106
	ITEM12	2.679E-02	-.113	.158	-1.038E-02
	ITEM13	-4.825E-02	-.258	-8.706E-03	9.484E-03
	ITEM14	.109	-.375	3.849E-02	-5.480E-02
	ITEM15	-6.321E-02	-7.588E-02	-4.007E-02	-1.040E-02
	ITEM16	-5.192E-02	-.199	-.148	-.139
	ITEM17	-6.273E-02	.151	-.256	9.572E-03
	ITEM18	-4.044E-03	-3.007E-02	.230	.164
	ITEM19	-.358	9.503E-02	-.306	-.390
	ITEM20	.681 <sup>a</sup>	-6.094E-02	-7.159E-03	.173
	ITEM27	-6.094E-02	.763 <sup>a</sup>	-.128	-.115
	ITEM28	-7.159E-03	-.128	.597 <sup>a</sup>	.101
	ITEM29	.173	-.115	.101	.647 <sup>a</sup>

a. Measures of Sampling Adequacy (MSA)

### Communalities

	Initial	Extraction
ITEM1	1.000	.619
ITEM2	1.000	.633
ITEM6	1.000	.485
ITEM11	1.000	.588
ITEM12	1.000	.642
ITEM13	1.000	.650
ITEM14	1.000	.616
ITEM15	1.000	.639
ITEM16	1.000	.514
ITEM17	1.000	.674
ITEM18	1.000	.717
ITEM19	1.000	.766
ITEM20	1.000	.537
ITEM27	1.000	.585
ITEM28	1.000	.690
ITEM29	1.000	.450

Extraction Method: Principal Component Analysis

### Total Variance Explained

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	3.905	24.405	24.405	3.905	24.405	24.405
2	2.110	13.188	37.593	2.110	13.188	37.593
3	1.598	9.991	47.584	1.598	9.991	47.584
4	1.118	6.987	54.571	1.118	6.987	54.571
5	1.072	6.697	61.268	1.072	6.697	61.268
6	.969	6.057	67.325			
7	.912	5.698	73.023			
8	.722	4.512	77.535			
9	.651	4.066	81.601			
10	.570	3.562	85.164			
11	.545	3.407	88.570			
12	.503	3.144	91.715			
13	.441	2.758	94.473			
14	.347	2.170	96.644			
15	.291	1.820	98.463			
16	.246	1.537	100.000			

Extraction Method: Principal Component Analysis

**Component Matrix<sup>a</sup>**

	Component				
	1	2	3	4	5
ITEM1	.301	-.352	1.465E-03	.190	.607
ITEM2	.613	-.164	.291	.257	-.283
ITEM6	.177	.583	-8.478E-02	.266	.189
ITEM11	.312	.562	-.362	.211	3.314E-04
ITEM12	.614	.359	-.360	2.823E-02	-7.117E-02
ITEM13	.608	-.158	-.103	.256	-.422
ITEM14	.669	-.365	-.181	-4.102E-02	1.001E-02
ITEM15	.485	.362	-.370	-.316	.189
ITEM16	.650	.161	-1.656E-02	-.240	8.762E-02
ITEM17	.532	-.528	-4.542E-02	-.301	.140
ITEM18	.327	-.599	-.191	.381	.264
ITEM19	.405	.239	.683	.168	.223
ITEM20	.411	.104	.509	.203	-.238
ITEM27	.664	-.108	-.170	-.167	-.277
ITEM28	.354	6.431E-02	.474	-.579	3.394E-02
ITEM29	.412	.372	.219	8.382E-02	.293

Extraction Method: Principal Component Analysis

a. 5 components extracted.

**Rotated Component Matrix<sup>a</sup>**

	Component				
	1	2	3	4	5
ITEM1	-6.279E-02	3.005E-02	.144	.767	7.428E-02
ITEM2	.651	-2.913E-02	.437	.124	3.941E-02
ITEM6	-.115	.620	.247	-4.813E-02	-.156
ITEM11	.126	.743	8.361E-03	-9.385E-02	-.107
ITEM12	.391	.657	-7.850E-03	1.519E-02	.182
ITEM13	.786	.132	8.482E-02	8.577E-02	-2.824E-02
ITEM14	.549	9.182E-02	-4.229E-02	.446	.325
ITEM15	6.407E-02	.632	-.121	5.316E-02	.467
ITEM16	.271	.374	.186	.115	.503
ITEM17	.351	-.159	-7.316E-02	.485	.533
ITEM18	.315	-9.196E-02	-6.924E-02	.763	-.148
ITEM19	-7.176E-04	8.733E-02	.855	9.111E-02	.134
ITEM20	.374	-2.102E-02	.619	-.111	1.895E-02
ITEM27	.632	.189	-4.171E-02	6.407E-02	.380
ITEM28	1.241E-02	-.118	.373	-.140	.719
ITEM29	-4.305E-02	.412	.487	.117	.166

Extraction Method: Principal Component Analysis.

Rotation Method: Varimax with Kaiser Normalization.

a. Rotation converged in 17 iterations.

### Component Transformation Matrix

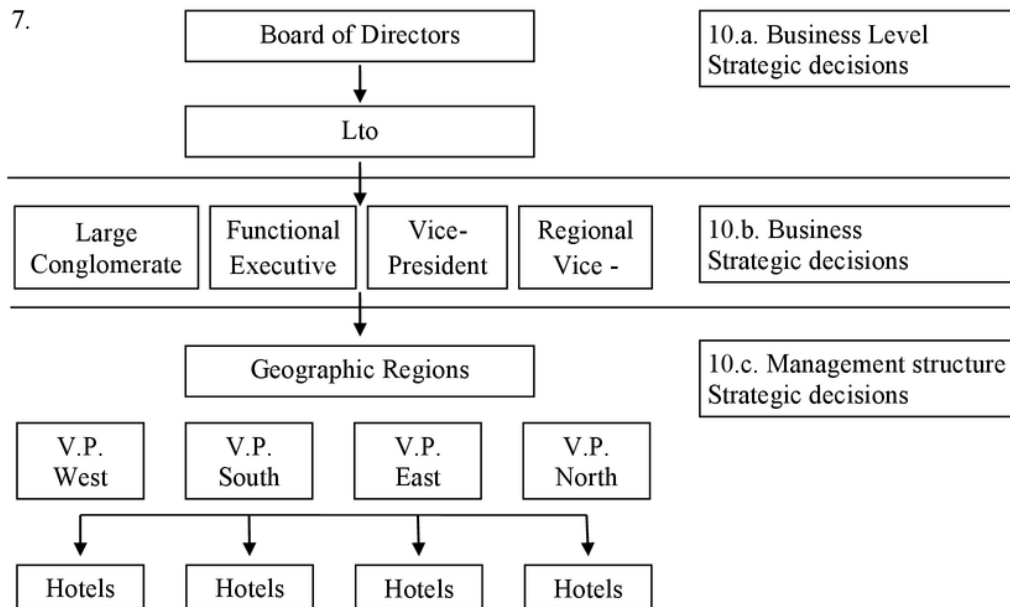
Component	1	2	3	4	5
1	.650	.411	.329	.325	.441
2	-.292	.727	.270	-.559	-.035
3	-.121	-.480	.847	-.159	.112
4	.206	.156	.288	.289	-.876
5	-.659	.220	.140	.688	.157

Extraction Method: Principal Component Analysis.

Rotation Method: Varimax with Kaiser Normalization.

## Appendix 16: Sample Student's Work on Piloting Stage

1. Hotels, both large and small are exciting place
2. The writer related it with his/her experience or his/her explanation.
3. Hotels
4. Managers must know how to work with groups
5. Working with groups on-the-job performance of work tasks
6. When Manager supervising employees on the job and when providing training to staff members



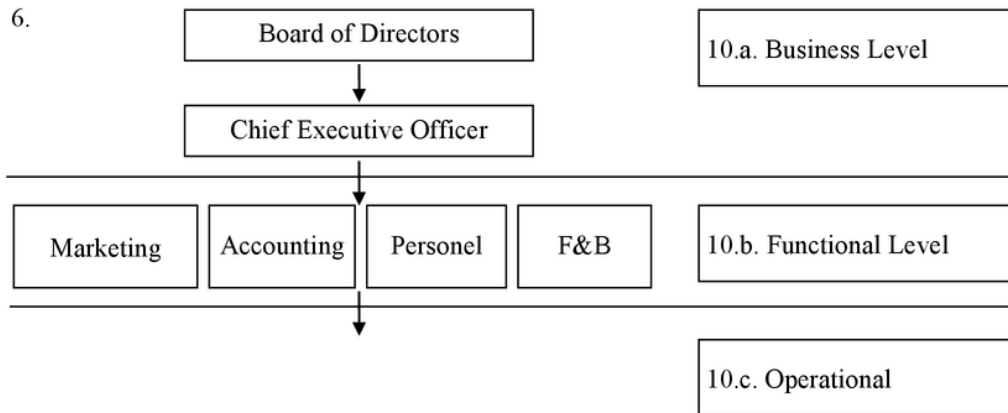
8. Japanese noodles
9. Both *soba* and *udon* are Japanese noodles and vary in the taste and thickness recording to the region in which they are made.
10. People in the east favor *soba* noodles, which are made from buckwheat flour kneaded in to five brown threads, white those in the west prefer thick wheat *udon* noodles
11. It served in a very simple, hot soy - sauce based broth or plain boiled to be dipped into liquid condiment.
12. The foreign influence in Japanese culinary habits in very much in evidence, with curry *soba* and curry *udon* being among the most popular Japanese noodles variants
13. Soba
14. Because it is just one of the countless Japanese food that have not made journey overseas.
15. General Manager (GM)
16. To discovered that the management and staff of these hotels felt that their hotel was by for the best in the city.
17. This feeling of being the best was widespread among the management and stuff and on mistakable in each hotel.

18. Because according to the philosophy of self - concept, or a culture whatever it is called, it's quite clear that developing the right kind of culture is an important goal for all well managed hotels.
19. Eight of ten general managerd (GMs) that were studied feel that understanding people is one of their main strengths.
20. Some seem to be natural when it comes to people.
21. They considered their ability to develop people as the major reasons for their success.
22. I agree with him because a quiet person will do his job/work with carefully.
23. *Karihara's* book
24. The cooker / Japanese cooker
25. Because since being set up in 2004, the government is intellectual property strategy headquarters his followed a policy trying to raise the brand value and push global saled of Japanese food.
26. Good reputation
27. No, it will not
28. No, it hasn't
29. They can establish Japanese food like a brand what Italian, Mexican, and Indian food has.
30. It influence in come of Japanese food.
31. Only by increasing opportunities for people to sample genuine Japanese cooking
32. Because they have a brand



## Appendix 17: Sample Student's Work on Test Tryout

1. Hotels are exciting places
2. Hotel are exciting
3. Supervising and training
4. Main steps
5. In morning briefing, manager give staff's job description, and accept all new idea to develop the hotel
- 6.



7. Japanese noodles with haven't made journey overseas
8. They are noodles *soba* are made from buckwheat and *udon* are from wheat
9. Because
10. Dipped into hot soy sauce based broth
11. Curry
12. *Soba* noodles
13. Because this kind of food haven't made journey overseas
14. To research the GM
15. Feeling of being the best
16. To reach the important good
17. First, third, fifth, seventh, eight
18. Second, fourth, sixth
19. Understanding elegance cleverness
20. Not really, because give their ideas straightly to their staffs. A lot of quiet persons are working according to their own thinking
21. One of Japanese woster cooking
22. Because the brand value of Japanese food is increasing
23. Push global sales raise the brand value of Japanese food
24. No, they won't
25. Yes, it has
26. Good reputation
27. Only by increasing opportunities for people to sample genuine Japanese food
28. Because they had promote their food overseas
- 29.

30. No, they don't, because the goodness only in one side, and the other side is not pretty good

31. In Reception & lobby

# ESP IN PRACTICE

---

## ORIGINALITY REPORT

---

13%

SIMILARITY INDEX

12%

INTERNET SOURCES

3%

PUBLICATIONS

9%

STUDENT PAPERS

---

## PRIMARY SOURCES

---

1

[dosfan.lib.uic.edu](http://dosfan.lib.uic.edu)

Internet Source

1%

2

[teaching-english-to-you.blogspot.com](http://teaching-english-to-you.blogspot.com)

Internet Source

1%

3

[www.readbag.com](http://www.readbag.com)

Internet Source

1%

4

Submitted to Macquarie University

Student Paper

<1%

5

[blogrizkirmd.blogspot.com.au](http://blogrizkirmd.blogspot.com.au)

Internet Source

<1%

6

[saomaidata.org](http://saomaidata.org)

Internet Source

<1%

7

Submitted to Republic of the Maldives

Student Paper

<1%

8

Submitted to La Trobe University

Student Paper

<1%

9

[www.ukessays.com](http://www.ukessays.com)

Internet Source

<1%

---

10

Submitted to Queen's University of Belfast

Student Paper

&lt;1 %

11

Submitted to University of Johannesburg

Student Paper

&lt;1 %

12

[www.slideshare.net](http://www.slideshare.net)

Internet Source

&lt;1 %

13

148.216.10.92

Internet Source

&lt;1 %

14

[repository.uinjkt.ac.id](http://repository.uinjkt.ac.id)

Internet Source

&lt;1 %

15

[ira.le.ac.uk](http://ira.le.ac.uk)

Internet Source

&lt;1 %

16

[perlinguam.journals.ac.za](http://perlinguam.journals.ac.za)

Internet Source

&lt;1 %

17

[unmas-library.ac.id](http://unmas-library.ac.id)

Internet Source

&lt;1 %

18

[scholar.lib.vt.edu](http://scholar.lib.vt.edu)

Internet Source

&lt;1 %

19

[scholar.sun.ac.za](http://scholar.sun.ac.za)

Internet Source

&lt;1 %

20

[rezkifirdaussps.blogspot.com](http://rezkifirdaussps.blogspot.com)

Internet Source

&lt;1 %

21

Submitted to Kaplan College

Student Paper

&lt;1 %

22	<a href="https://assets.cambridge.org">assets.cambridge.org</a> Internet Source	<1 %
23	<a href="https://www.docstoc.com">www.docstoc.com</a> Internet Source	<1 %
24	Submitted to Grand Canyon University Student Paper	<1 %
25	<a href="https://ftp.vobs.at">ftp.vobs.at</a> Internet Source	<1 %
26	<a href="https://eprints.qut.edu.au">eprints.qut.edu.au</a> Internet Source	<1 %
27	Submitted to University of South Florida Student Paper	<1 %
28	Submitted to University of Westminster Student Paper	<1 %
29	<a href="https://anatomiestreetsoldier.wordpress.com">anatomiestreetsoldier.wordpress.com</a> Internet Source	<1 %
30	Ronald Jay Cohen. Psychological Testing and Assessment, 2005 Publication	<1 %
31	<a href="https://journal.teflin.org">journal.teflin.org</a> Internet Source	<1 %
32	Submitted to Higher Education Commission Pakistan Student Paper	<1 %

33	<a href="https://dspace.lboro.ac.uk">dspace.lboro.ac.uk</a> Internet Source	<1 %
34	<a href="http://www.englishaustralia.com.au">www.englishaustralia.com.au</a> Internet Source	<1 %
35	Submitted to Bahcesehir University Student Paper	<1 %
36	<a href="http://koreatesol.org">koreatesol.org</a> Internet Source	<1 %
37	<a href="http://perpus.stkipkusumanegara.ac.id">perpus.stkipkusumanegara.ac.id</a> Internet Source	<1 %
38	Submitted to University of Southern California Student Paper	<1 %
39	Submitted to Walden University Student Paper	<1 %
40	Submitted to University of Queensland Student Paper	<1 %
41	Submitted to American Public University System Student Paper	<1 %
42	<a href="https://espace.curtin.edu.au">espace.curtin.edu.au</a> Internet Source	<1 %
43	Submitted to University of Reading Student Paper	<1 %

44	Submitted to University of Houston System Student Paper	<1 %
45	<a href="http://www.egris.ch">www.egris.ch</a> Internet Source	<1 %
46	Submitted to Chulalongkorn University Student Paper	<1 %
47	Submitted to University of Glamorgan Student Paper	<1 %
48	<a href="http://upit.ro">upit.ro</a> Internet Source	<1 %
49	"19. Language for specific purposes vs. general language", Walter de Gruyter GmbH, 2009 Publication	<1 %
50	Submitted to University of Wales, Bangor Student Paper	<1 %
51	<a href="http://xminxx.cmjy.sh.cn">xminxx.cmjy.sh.cn</a> Internet Source	<1 %
52	Submitted to University of Nottingham Student Paper	<1 %
53	Springer Theses, 2013. Publication	<1 %
54	<a href="http://www.spbea.org.fj">www.spbea.org.fj</a> Internet Source	<1 %

55	<a href="http://eprints.undip.ac.id">eprints.undip.ac.id</a> Internet Source	<1 %
56	Submitted to University of Leeds Student Paper	<1 %
57	<a href="http://www.aepc.gov.np">www.aepc.gov.np</a> Internet Source	<1 %
58	Submitted to University of Exeter Student Paper	<1 %
59	Cecil R. Reynolds. "The fallacy of "two years below grade level for age" as a diagnostic criterion for reading disorders", Journal of School Psychology, 1981 Publication	<1 %
60	Submitted to Universiti Utara Malaysia Student Paper	<1 %
61	Submitted to University of Edinburgh Student Paper	<1 %
62	Submitted to Aston University Student Paper	<1 %
63	<a href="http://www.pierpont.edu">www.pierpont.edu</a> Internet Source	<1 %
64	Submitted to 5027 Student Paper	<1 %



65

Internet Source

&lt;1 %

66

[theses.ucalgary.ca](https://theses.ucalgary.ca)

Internet Source

&lt;1 %

67

[www.rug.nl](http://www.rug.nl)

Internet Source

&lt;1 %

68

[www.academypublication.com](http://www.academypublication.com)

Internet Source

&lt;1 %

69

[www.klis.pf.ukf.sk](http://www.klis.pf.ukf.sk)

Internet Source

&lt;1 %

70

[tailieu.vn](http://tailieu.vn)

Internet Source

&lt;1 %

71

[www.imla.co](http://www.imla.co)

Internet Source

&lt;1 %

72

[www.ppk.elte.hu](http://www.ppk.elte.hu)

Internet Source

&lt;1 %

73

[www.canada.metropolis.net](http://www.canada.metropolis.net)

Internet Source

&lt;1 %

74

[files.eric.ed.gov](http://files.eric.ed.gov)

Internet Source

&lt;1 %

75

[ocean.otr.usm.edu](http://ocean.otr.usm.edu)

Internet Source

&lt;1 %

76

[thesis.swu.ac.th](http://thesis.swu.ac.th)

Internet Source

&lt;1 %

77	<a href="http://www.isqae.com">www.isqae.com</a> Internet Source	<1 %
78	<a href="http://www.auburn.edu">www.auburn.edu</a> Internet Source	<1 %
79	Submitted to CVC Nigeria Consortium Student Paper	<1 %
80	Submitted to University of Salford Student Paper	<1 %
81	<a href="http://www.geocities.com">www.geocities.com</a> Internet Source	<1 %
82	Olivo, S. A., L. G. Macedo, I. C. Gadotti, J. Fuentes, T. Stanton, and D. J Magee. "Scales to Assess the Quality of Randomized Controlled Trials: A Systematic Review", Physical Therapy, 2008. Publication	<1 %
83	<a href="http://library.iugaza.edu.ps">library.iugaza.edu.ps</a> Internet Source	<1 %
84	<a href="http://ddms.usim.edu.my">ddms.usim.edu.my</a> Internet Source	<1 %
85	<a href="http://kuscholarworks.ku.edu">kuscholarworks.ku.edu</a> Internet Source	<1 %
86	<a href="http://link.springer.com">link.springer.com</a> Internet Source	<1 %

87	<a href="https://dspace.bilkent.edu.tr">dspace.bilkent.edu.tr</a> Internet Source	<1 %
88	<a href="http://www.ecrif.com">www.ecrif.com</a> Internet Source	<1 %
89	<a href="http://www.atgci.org">www.atgci.org</a> Internet Source	<1 %
90	<a href="http://uir.unisa.ac.za">uir.unisa.ac.za</a> Internet Source	<1 %
91	<a href="http://www.ncpublicschools.org">www.ncpublicschools.org</a> Internet Source	<1 %
92	<a href="http://eprints.hud.ac.uk">eprints.hud.ac.uk</a> Internet Source	<1 %
93	<a href="http://eprints.utar.edu.my">eprints.utar.edu.my</a> Internet Source	<1 %
94	<a href="http://www.solki.jyu.fi">www.solki.jyu.fi</a> Internet Source	<1 %
95	<a href="http://www.scribd.com">www.scribd.com</a> Internet Source	<1 %
96	<a href="http://core.ac.uk">core.ac.uk</a> Internet Source	<1 %
97	<a href="http://libres.uncg.edu">libres.uncg.edu</a> Internet Source	<1 %
98	<a href="http://lib.unnes.ac.id">lib.unnes.ac.id</a> Internet Source	<1 %

99	<a href="http://eprints.usq.edu.au">eprints.usq.edu.au</a> Internet Source	<1 %
100	<a href="http://digi.library.tu.ac.th">digi.library.tu.ac.th</a> Internet Source	<1 %
101	<a href="http://eprints.uny.ac.id">eprints.uny.ac.id</a> Internet Source	<1 %
102	<a href="http://keera.or.kr">keera.or.kr</a> Internet Source	<1 %
103	"Proceedings of the Twelfth International Conference on Management Science and Engineering Management", Springer Nature, 2019 Publication	<1 %
104	<a href="http://digitalcommons.fiu.edu">digitalcommons.fiu.edu</a> Internet Source	<1 %
105	<a href="http://repo.iain-tulungagung.ac.id">repo.iain-tulungagung.ac.id</a> Internet Source	<1 %
106	Submitted to University of Auckland Student Paper	<1 %
107	<a href="http://jurnal.unimed.ac.id">jurnal.unimed.ac.id</a> Internet Source	<1 %
108	<a href="http://jlt-polinema.org">jlt-polinema.org</a> Internet Source	<1 %

Mebarki, Zahia. "Factors Underlying the

109	Reading Performance of Algerian Microbiology Students", Procedia - Social and Behavioral Sciences, 2011.	<1 %
Publication		

---

110	dspace.unach.edu.ec	<1 %
Internet Source		

---

111	Submitted to Capital Education	<1 %
Student Paper		

---

112	www.academypublisher.com	<1 %
Internet Source		

---

113	www.ielts.org	<1 %
Internet Source		

---

114	bahasainggrisunpar.blogspot.com	<1 %
Internet Source		

---

115	www.ohiolink.edu	<1 %
Internet Source		

---

Exclude quotes      On

Exclude matches      Off

Exclude bibliography      On