COMPARING TWO TRANSLATION ASSESSMENT MODELS: CORRELATING STUDENT REVISIONS AND PERSPECTIVES

A dissertation submitted to Kent State University in partial fulfillment of the requirements for the degree of Doctor of Philosophy

by

Haru Deliana Dewi

December 2015
Dissertation written by

Haru Deliana Dewi

S.S., Universitas Indonesia, 1998

M.Hum., Universitas Indonesia, 2006

Ph.D., Kent State University, 2015

Approved by

Sue Ellen Wright, Chair, Doctoral Dissertation Committee

Richard Kelly Washbourne, Members, Doctoral Dissertation Committee

Erik Angelone

Brian Huot

Kevin Floyd

Accepted by

Keiran Dunne, Chair, Department of Modern and Classical Language Studies

James L. Blank, Dean, College of Arts and Sciences
# TABLE OF CONTENTS

TABLE OF CONTENTS........................................................................................................ iv

LIST OF FIGURES........................................................................................................... viii

LIST OF TABLES............................................................................................................. ix

ACKNOWLEDGMENTS........................................................................................................ xi

CHAPTER 1: INTRODUCTION................................................................................................. 1

1.1 Background and Overview of the Study................................................................. 2

1.2 Statement of the Problem & Objectives of the Study............................................ 4

1.3 Hypotheses.................................................................................................................. 6

1.4 Significance of the Study............................................................................................ 7

1.5 Overview of the Dissertation.................................................................................... 9

CHAPTER 2: DEFINING CRITICAL TERMS........................................................................ 11

2.1 Assessment and Evaluation...................................................................................... 11

2.2 Teacher-Centered Approach and Student-Centered Approach............................. 18

2.3 Assessment Types..................................................................................................... 22

2.3.1 Summative and Formative Assessment............................................................. 23

2.3.2 Feedback............................................................................................................... 29

2.3.3 Certification and Accreditation........................................................................... 33

2.3.4 Assessment Models.............................................................................................. 35

CHAPTER 3: TRANSLATION ASSESSMENT DEVELOPMENT...................................... 40

3.1 Translation Quality Assessment (TQA)................................................................... 40

3.2 Skopos Theory.......................................................................................................... 45

3.3 Theoretical and Descriptive Theory on Translation Assessment........................... 49

3.4 Empirical Research on Translation Assessment.................................................... 51
6.2.2 Weaknesses of the Pilot Projects and Changes for the Second Projects.......................................................... 103

6.3 Formal or Second Project Result Analysis and Discussion............. 104

6.3.1 Inter-Rater Reliability.................................................................... 105

6.3.2 The Quantitative Analysis and the Interpretation of the Second Projects.................................................................. 110

6.4 Online Survey Results........................................................................ 121

6.4.1 Pilot Project Survey Results..................................................... 122

6.4.2 Formal or Second Project Survey Results................................. 127

CHAPTER 7: CONCLUSIONS: SUMMARIES, LIMITATIONS, AND FUTURE RESEARCH......................................................... 132

7.1 Summaries..................................................................................... 132

7.1.1 The Relative Effectiveness of the Two Models....................... 133

7.1.2 Respondents’ Preferences on Translation Assessment Models....................................................................................... 134

7.1.3 Correlation between Respondent Perspectives and Their Translation Results.......................................................... 135

7.1.4 Translation Error Types and Two Assessment Models.......... 136

7.2 Limitations.................................................................................... 137

7.3 Future Research............................................................................. 140
# REFERENCES

APPENDICES

<table>
<thead>
<tr>
<th>Appendix</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>THE LBI BANDSCALE</td>
<td>157</td>
</tr>
<tr>
<td>B</td>
<td>THE ATA FRAMEWORK</td>
<td>159</td>
</tr>
<tr>
<td>C</td>
<td>FLYERS AND CONSENT FORMS</td>
<td>165</td>
</tr>
<tr>
<td>D</td>
<td>FIRST AND FINAL ASSIGNMENT INSTRUCTIONS</td>
<td>179</td>
</tr>
<tr>
<td>E</td>
<td>THE SURVEYS</td>
<td>191</td>
</tr>
<tr>
<td>F</td>
<td>ILR (THE INTERAGENCY LANGUAGE ROUNDTABLE) SCALES</td>
<td>195</td>
</tr>
<tr>
<td>G</td>
<td>IELTS (THE INTERNATIONAL ENGLISH LANGUAGE TESTING SYSTEM) SCALES</td>
<td>201</td>
</tr>
<tr>
<td>H</td>
<td>NAATI (THE NATIONAL ACCREDITATION AUTHORITY FOR TRANSLATORS AND INTERPRETERS)</td>
<td>203</td>
</tr>
</tbody>
</table>
# List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 6.1.</td>
<td>Frequency of Error Types (Indonesian-English)</td>
<td>93</td>
</tr>
<tr>
<td>Figure 6.2.</td>
<td>Model 1 for Sampling Distribution If $t_{\text{obt}}$ is beyond $t_{\text{crit}}$</td>
<td>100</td>
</tr>
<tr>
<td>Figure 6.3.</td>
<td>Model 2 for Sampling Distribution If $t_{\text{obt}}$ is not beyond $t_{\text{crit}}$</td>
<td>100</td>
</tr>
<tr>
<td>Figure 6.4.</td>
<td>Sampling Distribution of the Pilot Project Result Fall 2013</td>
<td>102</td>
</tr>
<tr>
<td>Figure 6.5.</td>
<td>Sampling Distribution of the ED1 of the Second Projects</td>
<td>111</td>
</tr>
<tr>
<td>Figure 6.6.</td>
<td>Sampling Distribution of the ED2 of the Second Projects</td>
<td>113</td>
</tr>
<tr>
<td>Figure 6.7.</td>
<td>Sampling Distribution of the Paragraph 1 ED1 of the Second Projects</td>
<td>115</td>
</tr>
<tr>
<td>Figure 6.8.</td>
<td>Sampling Distribution of the Paragraph 1 ED2 of the Second Projects</td>
<td>116</td>
</tr>
<tr>
<td>Figure 6.9.</td>
<td>Sampling Distribution of the Paragraph 2 ED1 of the Second Projects</td>
<td>118</td>
</tr>
<tr>
<td>Figure 6.10.</td>
<td>Sampling Distribution of the Paragraph 2 ED2 of the Second Projects</td>
<td>119</td>
</tr>
<tr>
<td>Figure 6.11.</td>
<td>How General or Specific the Assessment Models</td>
<td>122</td>
</tr>
<tr>
<td>Figure 6.12.</td>
<td></td>
<td>123</td>
</tr>
<tr>
<td></td>
<td>The ATA Framework: General or Specific?</td>
<td></td>
</tr>
<tr>
<td></td>
<td>The LBI Bandscale: General or Specific?</td>
<td></td>
</tr>
<tr>
<td>Figure 6.13.</td>
<td></td>
<td>124</td>
</tr>
<tr>
<td></td>
<td>The ATA Framework: Simple or Complicated?</td>
<td></td>
</tr>
<tr>
<td></td>
<td>The LBI Bandscale: Simple or Complicated?</td>
<td></td>
</tr>
<tr>
<td>Figure 6.14.</td>
<td>The Results of Questions 1 and 2</td>
<td>127</td>
</tr>
<tr>
<td>Figure 6.15.</td>
<td>The Results of Questions 3 and 4</td>
<td>128</td>
</tr>
<tr>
<td>Figure 6.16.</td>
<td>The Results of Questions 5 and 6</td>
<td>129</td>
</tr>
</tbody>
</table>
LIST OF TABLES

Table 2.1. Assessment Vs. Evaluation ................................................................. 14
Table 2.2. Differences between Assessment and Evaluation ............................. 14
Table 2.3. Differences between Processes of Assessment and Evaluation ............ 15
Table 2.4. Comparison of Evaluation and Assessment ......................................... 15
Table 6.1. The Canadian Language Quality Measurement System (SICAL) ............... 86
Table 6.2. Translation Error Typology of Lee & Ronowick (Korean to English) ....... 89
Table 6.3. Translation Errors based on This Study (Indonesian to English) .......... 92
Table 6.4. Pilot Project Data ................................................................................ 97
Table 6.5. T-Test Results of the Fall 2013 Pilot Project ....................................... 101
Table 6.6. Pearson’s r Correlation ..................................................................... 106
Table 6.7. Second Project Data .......................................................................... 107
Table 6.8. Correlation of All ED1s and ED2s for the ATA Framework and the LBI Bandscale ................................................................. 108
Table 6.9. Correlation of ED1 and ED2 for the ATA Framework ......................... 109
Table 6.10. Correlation of ED1 and ED2 for the LBI Bandscale ......................... 109
Table 6.11. The Paired Samples T-Test for ED1 for the ATA Framework and the LBI Bandscale ........................................................................ 111
Table 6.12. The Paired Samples T-Test for ED2 for the ATA Framework and the LBI Bandscale ........................................................................ 112
Table 6.13. Paragraph 1 Data ............................................................................. 113
Table 6.14. The Independent Samples T-Test for Paragraph 1 ED1 for the ATA Framework and the LBI Bandscale .................................................. 114
Table 6.15. The Independent Samples T-Test for Paragraph 1 ED2 for the ATA Framework and the LBI Bandscale........................................... 115

Table 6.16. Paragraph 2 Data........................................................................................................ 117

Table 6.17. The Independent Samples T-Test for Paragraph 2 ED1 for the ATA Framework and the LBI Bandscale........................................... 117

Table 6.18. The Independent Samples T-Test for Paragraph 2 ED2 for the ATA Framework and the LBI Bandscale........................................... 118

Table 6.19. Reasons Why the Assessment Models are Useful.......................... 126

Table 6.20. The Results of Question 7.......................................................... 130

Table 7.1. The Summary of Statistical Calculations......................................................... 133
ACKNOWLEDGMENTS

I would like to express my sincere gratitude to my advisor, Dr. Sue Ellen Wright, for her support, patience, and insightful and very helpful comments throughout the entire dissertation process. Her expertise, experience, and encouragement were very valuable during the time I spent in the doctoral program. I would also like to thank all my dissertation committee members, Dr. Richard Kelly Washbourne, Dr. Erik Angelone, Dr. Brian Huot, and Dr. Kevin Floyd for their comments, suggestions, recommendations, and challenging questions during my dissertation writing process and the defense. Thank you very much to the faculty of the Department of Modern and Classical Language Studies, the Institute for Applied Linguistics, and the English Department of KSU for their support during my studies. I am also very grateful to the English Study Program, the Faculty of Humanities, Universitas Indonesia and all my colleagues who always had my back and encouraged me to complete my doctoral program. Many thanks to my family in Jakarta, all my classmates, and friends for their love, friendship, and support. Finally, I would like to thank the staff of LBI FIB UI who provided assistance and the respondents who participated in this study.

Haru Deliana Dewi

Kent, Ohio

December 2015
CHAPTER 1
INTRODUCTION

Translation Studies and translator training in the US have evolved in parallel with translator training in Europe and Canada, and some practices reflect a long tradition that differs from practice in Indonesia. Furthermore, in translation assessment, these two countries apply different models. To be more specific, the Translation Center (PP LBI – Lembaga Bahasa Internasional) classes of Universitas Indonesia (UI) in Jakarta, Indonesia, use an assessment scale to assess student final translations, while several classes of the Kent State University (KSU) Translation Program use an error analysis or an analytic assessment to evaluate student draft and final translations. The translation students in LBI UI usually have feedback of their translations in the class discussion but rarely have the chance to receive feedback in writing or to prepare a second draft that translation trainers can use to gage their progress. In KSU translation classes, the students receive formative assessment from both each other in class and from their instructors, and that they often have the chance to correct their work. Because of these differences, it is interesting to compare the effectiveness of these two different models in assessing translations of the participants enrolling to translation classes in Indonesia to discover whether their effectiveness is similar or not. This study specifically attempts to reveal the effectiveness of these two models applied as formative assessment and to reveal student perspectives towards those two assessment models.
1.1 Background and Overview of the Study

In the 1990s and in the early 2000s, a few scholars in Translation Studies stated that there was a lack of empirical research on translation assessment, and that empirical research involving student perspectives on translation assessment was even scarcer. In 1992, Pym confirmed his suspicion that there was no empirical research readily available for citation on translation assessment (1992:279-280). Moreover, Hatim and Mason remarked that “the assessment of translator performance is an activity which, despite being widespread, is under-researched and under-discussed” (1997:197). Melis and Albir (2001) recommended scholars in Translation Studies do more research on assessment. They stated that translation assessment had hardly existed yet in this field of study, and when it was practised, it had not been carried out objectively and in a disciplined fashion (2001:273).

Even in the more recent publications, Waddington states that research in the field of translation quality assessment was mainly theoretical and descriptive before 2001 (2001a:16), and that there was “an almost complete absence of empirical studies” in the area of translation quality assessment (2001a:18). This means from the emergence of Translation Studies in the 1970s up to 2001, even though a number of scholars had been discussing translation evaluation/assessment, almost none of them was conducting empirical research on it. Waddington was among the very few who started conducting empirical research in translation assessment and published the results in 2001. From 2001 to the present time, there has still been only limited empirical research conducted in translation assessment. According to Arango-Keeth and Koby, out of four fields in Translation Studies, including theory, practice, pedagogy, and evaluation (or quality assessment), translation evaluation is still the least developed (2003:117). Furthermore, Angelelli and Jacobson maintain that there is still a lack of empirical research on assessing processes and products of translation in the fields of
industry, pedagogy, and research (2009:4). Conde also states that “evaluation is still a field in which much remains to be explored” (2012a:68).

In an effort to address this need, a pilot observational study has been conducted in order to discover the effectiveness of two translation assessment models. The first model is the Indonesian translation bandscale (henceforth: the LBI Bandscale), which is a translation assessment scale used in Pusat Penerjemahan Lembaga Bahasa Internasional Universitas Indonesia (PP LBI UI = the Translation Center of the International Language Institution of the University of Indonesia) in Jakarta, Indonesia. This model provides a general description of translation results from very good ones (Grade A) to very bad ones (Grade D). The description includes positive and negative feedback on a translation result from the standpoint of both source text comprehension and target text writing. The second model is the American Translators Association Framework for Error Marking (henceforth: the ATA Framework, which in this study is the one that has been modified to be used to assess formative exercises) used by the ATA to assess their certification examinations. This model provides detailed explanations of many types of errors and contains granular details of errors.

These two models were originally intended to be used for summative assessment, but have been applied here for formative feedback and for the evaluation of revisions of translations produced by the respondents. In other words, this study focuses on applying these two models to assess formative exercises. An in-depth explanation about assessment scales and error analysis, as well as summative and formative assessment, can be seen in Chapter 2, and the details of these two models will be extensively discussed in Chapter 4. The study will then correlate student perspectives on translation assessment, which can be seen in the answers to surveys conducted to student response to assessment (see Appendix E), with their translation results assessed by me (Rater 1) and by an independent judge (Rater 2) in the formal projects.
The participants involved were students enrolled in a General Translation Course held at Pusat Penerjemahan Lembaga Bahasa Internasional Universitas Indonesia (PP LBI UI – the Translation Center of the International Language Institution of the University of Indonesia) in Jakarta, Indonesia. The tasks given to them were independent from course assignments and were not allowed to influence their grades in class according to IRB rules. The study consisted of a pair of pilot projects and a set of follow-up (formal or second) projects. The students were invited to voluntarily join the projects. This was a remote observational study as it was conducted from Kent, OH, while the participants were in Jakarta, Indonesia. The initial pilot projects were conducted in the beginning of October 2013 and in Spring 2014, respectively. The follow-up projects were carried out in Summer 2014, Fall 2014, and Spring 2015. The details of the methodology are presented in Chapter 5.

1.2 Statement of the Problem and Objectives of the Study

As a result of the continuing dearth of empirical research conducted in translation evaluation or assessment, it is not easy to discover which translation assessment models are considered effective, especially for assessing formative exercises. The differences between assessment and evaluation will be discussed in Chapter 2, and the relative lack of empirical research focusing on formative translation assessment is explained in detail in Chapter 3. Furthermore, empirical research that attempts to reveal or include student perspectives is scarce. “Student perspectives” here refer to the opinions or points of view of the students, in this case the respondents of the study, on the two assessment models applied. This study includes student perspectives in an effort to contribute and initiate research in translation formative assessment and to reveal student perspectives on translation assessment. The perspectives will be revealed from the answers to two surveys. The surveys are online simple
surveys. The survey for the pilot projects is different from the survey for the follow-up projects, but the theme is the same. The detailed explanation of the surveys is presented in Chapter 5.

There are three research questions in this study. The first question asks which translation model is more effective in helping students improve their translations, the ATA Framework or the LBI Bandscale. Through this question, the study aims to discover the effectiveness of one or both models in the context of university translation classes. As noted above, both models were initially designed and used for summative assessment, and there have been several studies conducted to discover the effectiveness of these models for summative assessment (see Chapter 3). However, by applying them as formative assessment, this study represents an initial effort to establish best practices for developing a long-term research study on this theme. Moreover, in order to promote these two models as formative assessment models for translation classes in LBI UI Translation Center, this study can provide some insights concerning the effectiveness of these models in assessing students’ translations and revisions.

The second question asks which translation assessment model students prefer for improving their work. The surveys are designed to answer this question. With this question, the study attempts to reveal student perspectives on the two assessment models, and the question has several objectives. First, it aims to discover which translation model is more general or more specific and which one is more simple or more complicated according to the students. The surveys (the pilot project survey and the second project survey) include these questions. Second, this question attempts to find out which one is more effective based on students’ opinions. Third, the objective is to reveal which model the respondents (students) prefer for improving their work.
The third or last question asks whether the student perspectives revealed through the results of the surveys correlate with improvement in their translation results based on the revisions after having received assessment and feedback. Their revisions have been evaluated by Rater 1 (me) and Rater 2 (an independent judge) to ensure more objective results. As Rater 1, the person has to master both English and Indonesian well, has to know the two assessment models well, and has to be a translation academician (working at a university teaching translation to undergraduate and graduate students, and in this case working for the university under which the LBI translation classes are conducted) and a translator. It was only me who can fulfill those criteria as Rater 1 for this study. With this question, the study aims to provide new insights on the reaction of students and their actual performance related to translation assessment.

1.3 Hypotheses

Based on the research questions described above, this study poses three hypotheses:

1) There is a significant difference of effectiveness between the ATA Framework and the LBI Bandscale. These two assessment models are very different from each other, so it is expected that their effectiveness will be different as well. There will be statistical calculations to show whether this hypothesis can be supported or rejected (see Chapter 6), as this is based on quantitative results. The null hypothesis will be that there is no significant difference between the ATA Framework and the LBI Bandscale.

2) Relating to the first hypothesis, the second hypothesis is that students will prefer the ATA Framework to the LBI Bandscale in improving their translations, as indicated by the answers to the surveys. There will be no statistical calculations to discover
whether the second hypothesis can be supported or rejected, as this is not based on quantitative results. This hypothesis is based on qualitative results from the analysis of the survey results. The null hypothesis in this case is that the students (the respondents) will think that both assessment models are equally effective.

3) For the third hypothesis, it is predicted that there will be a positive correlation between student perspectives on the assessment models and their translation results, based on the improvement made in the revisions and the comparison of survey results. This hypothesis is based on qualitative results, not quantitative results, so there will not be any statistical calculations. A positive correlation between student perspectives and their translation results will mean that if students have positive perspectives on one or both assessment models, their translation results will also show improvement in the translation revisions. If the perspectives are positive, but there is no improvement, it means there is no correlation. If the perspectives are negative, but there is improvement, it also means there is no correlation. If the perspectives are negative, and there is no improvement, it means there is a negative correlation. What is expected from this hypothesis is that there will be a positive correlation rather than a negative one. The null hypothesis will be that there is no correlation between the perspectives and the translation results.

1.4 Significance of the Study

Given the cited lack of empirical research and of studies revealing student perspectives, this study is designed to contribute to the improvement of translation assessment provided to students. In addition, the study is expected to uncover student perspectives on translation assessment. This study focuses on formative exercises that are
expected to be applied in a student/learner-oriented approach classroom, and the opinions of students might help improve the exercises, although students may not always understand what is good for them. Besides that, students are one of the stakeholders in translation education. Kramer et al. state that “without stakeholders’ engagement in the assessment, the system has little chance of success” (2010:33). Hence, it will be interesting to discover what translation students actually think about translation assessment provided to them and to figure out the correlation between their perspectives and their translation results.

Furthermore, even if individual instructors claim that they have good assessment strategies which are compatible with students’ expectations and are encouraging students to improve their skills and knowledge, there is little evidence to support these beliefs if they have never conducted studies to determine the effectiveness of the assessment they apply. Kramer et al. advocate creating a culture of evidence to support the aims and claims on assessments of an institution, and as a body of knowledge to inform the stakeholders, including students (2010:45). Thus, in a culture of evidence, this study becomes significant as a tool to demonstrate whether one style of assessment applied really helps students learn and gain the skills and knowledge they expect.

In addition, it has often been said that there is a significant gap between translation theory and translation practice (Snell-Hornby et al. 1994:105), and this still also applies in translation assessment. Lauscher remarks that “scholarly approaches to translation evaluation have not yet been able to provide help for practical quality assessment because they do not account for the reality of translating and translations” (2000:158). Besides that, there is relatively little guidance for individual evaluators on how they might go about some sort of systematic assessment strategy. Thus, this observational study will also contribute to building a bridge between translation theory and translation practice, especially in translation assessment, as student perspectives could become one of the reality factors used to examine
whether the assessment given to them is absolutely effective for their improvement. Li states that assessment should be carried out to really fulfill the needs of both students and professional translators so that we can be sure that translation programs will always meet “the constantly changing social needs and the needs of the students” (2002:513).

1.5 Overview of the Dissertation

This dissertation consists of seven chapters. This first chapter has focused on the background, the research questions and hypotheses, the objectives and the significance of this study, and the overview of the dissertation. Chapter 2 provides definitions of a large number of critical terms which are important for this study. These terms must be understood first since there are several scholars from outside the U.S. who utilize terms which are not applied in this country. Then Chapter 3 covers the development of translation assessment models viewed from three major perspectives. The first part will explain about House’s Translation Quality Assessment (TQA), Skopos theory, and the influence of those theories in translation assessment. The second part will discuss the publications containing descriptive and theoretical concepts of translation evaluation/assessment. The third part will consist of the discussion of the publications on translation assessment/evaluation applying empirical research.

Chapter 4 will describe the projects conducted for this study. In this chapter, there will be detailed discussions of the models used in the projects, the LBI Bandscale and the ATA Framework. There will be also a subchapter discussing the similarities and the differences between these projects and the projects conducted by other scholars discussed in Chapter 3. Chapter 5 will discuss the methodology used for this study. All components of the projects will be elaborated, such as the participants, the materials, the procedures, and the surveys.
Chapter 6 presents an explanation of translation error definitions and error typology, and an analysis of the results of the study. It discusses both the pilot and the primary or formal projects. It also covers the online survey result discussion. Finally, Chapter 7 states the conclusions of the study, including the limitations and future research. This last chapter will provide several recommendations to continue research in the future in a similar theme.
CHAPTER 2
DEFINING CRITICAL TERMS

This chapter defines a number of terms that are used to describe translation assessment development in Chapter 3. All of the terms discussed are treated as a set of parallel concepts. The first terms are assessment and evaluation applied in education and in Translation Studies, and their differences will be elaborated. Other terms are the teacher-centered approach and the student-centered approach, and the contrast between them will be discussed. Then the description of the assessment types, ranging from summative and formative assessment in education and Translation Studies to certification and accreditation, will show similar and different concepts from one scholar to another, from one country to another. Definitions of feedback will also be explained as an important part of formative assessment. The last terms discussed present common assessment models used for writing and translation.

2.1 Assessment and Evaluation

In education in general, several scholars have highlighted differences between assessment and evaluation related to program assessment and evaluation and individual student assessment and evaluation. First, Angelo and Cross in defining program assessment and evaluation write that the focus of assessment is on learning, teaching, and outcomes, and
that assessment gives information for improving learning and teaching (1993:427). They affirm that “assessment is an interactive process between students and faculty that informs faculty how well their students are learning what they are teaching.” The information obtained from assessment is usually shared with students so that they can improve their learning and study habits, and this information is “learner-centered, course based, frequently anonymous, and not graded.” Who is usually anonymous is the students or participants involved in the assessment, while the teacher(s) and/or the person conducting the assessment is not anonymous. Evaluation, on the contrary, focuses on grades according to Angelo and Cross. It reflects more classroom components in addition to course content and mastery level and may include participation in discussion, cooperation, attendance, and verbal ability (Ibid.).

Palomba and Banta (1999) present a definition of program assessment that is widely and commonly used. However, they do not define what evaluation is and they do not provide any explanation of the differences between assessment and evaluation. According to them, in education, assessment describes the measurement of what an individual knows and can do (1999:1). They focus on the definition of assessment in student learning. In the complete version of their assessment definition, they state that “assessment is the systematic collection, review, and use of information about educational programs undertaken for the purpose of improving student learning and development” (4). In addition, they claim that assessment is more than the collection of data, so in order to make assessment work, educators have to be purposeful about what they collect (9).

According to Astin and Antonio, the term [program] *assessment* refers to two different activities. The first one is the gathering of information, which is called *measurement*, and the information is related to the *functioning* of students, staff, and institutions of higher education (2012:3). *Functioning* refers to the broad social purposes of a
college or university. The authors further explain that “the information may or may not be in numerical form, but the basic motive for gathering it is to improve the functioning of the institution and its people.” The second activity is the utilization of the information gathered for institutional and individual improvement, which is called evaluation. According to the authors, the term evaluation contains motivation and “the rendering of value judgments.”

Astin and Antonio provide a good example to differentiate measurement from evaluation, as an examination in a college course is considered as measurement, and evaluation refers to the ways the results are used (Ibid.). The following scholars will discuss the definition of individual assessment and evaluation.

Straight discusses the differences between assessment and evaluation as follows (2002:18-24). First of all, in terms of timing, assessment is formative, as it is meant to improve learning, while evaluation is summative because it is conducted in the final phase to gauge quality. Second, in terms of the measurement focus, assessment is process-oriented, as it reveals how learning is proceeding, whereas evaluation is product-oriented since it reveals what has been learned. Third, in terms of the relationship between administrator and recipient, assessment is reflective, as it internally defines criteria or goals. However, evaluation is prescriptive because it imposes external standards. Fourth, in terms of findings and uses, assessment is diagnostic as it identifies areas for improvement, whereas evaluation is “judgmental” since it arrives at an overall grade or score.

Fifth, in terms of ongoing modifiability of criteria and measures, assessment is flexible because it adjusts as problems are clarified, but evaluation is fixed since it rewards success and punishes failure. Sixth, in terms of standards of measurement, assessment is absolute, since it strives for ideal outcomes, while evaluation is comparative as it divides “better” students from “worse” students. Finally, seventh, in terms of relations between objects of A/E (Assessment/Evaluation), assessment is cooperative as students and/or
teachers learn from each other, whereas evaluation is competitive because students and/or teachers beat each other out (*Ibid.*). Tables 1 and 2 below clearly show the differences between assessment and evaluation according to Straight.

**Table 2.1: Assessment Vs Evaluation**

<table>
<thead>
<tr>
<th></th>
<th>Assessment</th>
<th>Evaluation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Formative:</strong></td>
<td>Ongoing to improve learning</td>
<td>Summative: Final to gauge quality</td>
</tr>
<tr>
<td><strong>Process-Oriented:</strong></td>
<td>How learning is going</td>
<td>Product-Oriented: What’s been learned</td>
</tr>
<tr>
<td><strong>Reflective:</strong></td>
<td>Internally defined criteria/goals</td>
<td>Prescriptive: Externally imposed standards</td>
</tr>
<tr>
<td><strong>Diagnostic:</strong></td>
<td>Identify areas for improvement</td>
<td>Judgmental: Arrive at an overall grade/score</td>
</tr>
<tr>
<td><strong>Flexible:</strong></td>
<td>Adjust as problems are clarified</td>
<td>Fixed: Reward success, punish failure</td>
</tr>
<tr>
<td><strong>Absolute:</strong></td>
<td>Strive for ideal outcomes</td>
<td>Comparative: Divide better from worse</td>
</tr>
<tr>
<td><strong>Cooperative:</strong></td>
<td>Learn from each other</td>
<td>Competitive: Beat each other out</td>
</tr>
</tbody>
</table>

Source: Straight, 2002:18-23

**Table 2.2: Differences between Assessment and Evaluation**

<table>
<thead>
<tr>
<th>Dimension of Difference</th>
<th>Assessment</th>
<th>Evaluation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Timing</td>
<td>Formative</td>
<td>Summative</td>
</tr>
<tr>
<td>Focus of measurement</td>
<td>Process-oriented</td>
<td>Product-oriented</td>
</tr>
<tr>
<td>Relationship between administrator and recipient</td>
<td>Reflective</td>
<td>Prescriptive</td>
</tr>
<tr>
<td>Findings, uses thereof</td>
<td>Diagnostic</td>
<td>Judgmental</td>
</tr>
<tr>
<td>Ongoing modifiability of criteria, measures thereof</td>
<td>Flexible</td>
<td>Fixed</td>
</tr>
<tr>
<td>Standards of measurement</td>
<td>Absolute</td>
<td>Comparative</td>
</tr>
<tr>
<td>Relation between objects of A/E</td>
<td>Cooperative</td>
<td>Competitive</td>
</tr>
</tbody>
</table>

Source: Straight, 2002:24

Baehr also distinguishes between assessment and evaluation. He states that assessment provides feedback on knowledge, skills, attitudes, and work products for the purpose of improving future performance and learning outcomes, while evaluation determines the level of quality of a performance or outcome and enables decision-making based on the level of quality demonstrated (2010:441). In addition, he explains that
assessment is the term used when looking at how the level of quality of a performance or outcome could be improved in the future, whereas evaluation is the term used to describe the determination of the level of quality (Ibid.). More explanations of the differences between assessment and evaluation by Baehr can be seen in the two tables below.

### Table 2.3: Differences between Processes of Assessment and Evaluation

<table>
<thead>
<tr>
<th></th>
<th>Assessment</th>
<th>Evaluation</th>
</tr>
</thead>
<tbody>
<tr>
<td>What is the purpose?</td>
<td>To improve the quality of future performances</td>
<td>To determine the quality of the present performance</td>
</tr>
<tr>
<td>On what is feedback based?</td>
<td>Observations; strongest and weakest points</td>
<td>Level of quality based on a set standard</td>
</tr>
<tr>
<td>What is included in the report?</td>
<td>What made the quality of the performance strong; how might one improve future performances</td>
<td>The quality of the performance, often compared to set standards</td>
</tr>
<tr>
<td>How is the report used?</td>
<td>To improve performance</td>
<td>To make judgments</td>
</tr>
</tbody>
</table>

Source: Baehr, 2010:442

### Table 2.4: Comparison of Evaluation and Assessment

<table>
<thead>
<tr>
<th></th>
<th>Assessment</th>
<th>Evaluation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Assessment provides feedback on performance; strengths, areas for improvement, and insights.</td>
<td>Evaluation determines whether a standard was met; success or failure.</td>
<td></td>
</tr>
<tr>
<td>Assessor provides feedback about how the performance can be improved.</td>
<td>Evaluator, using a higher set standard, determines that the measured performance was a failure.</td>
<td></td>
</tr>
<tr>
<td>Assessor provides feedback on what made the performance as strong as it was.</td>
<td>Evaluator, using a lower set standard, determines that the measured performance was a success.</td>
<td></td>
</tr>
</tbody>
</table>

Source: Baehr, 2010:443

In another discussion examining the differences between assessment and evaluation, Taras views differences based on the perspectives of the UK and much of the Anglophone world. She maintains that evaluation covers the macro spectrum, such as at the university, program, or course level, while assessment covers the micro level, such as the assignment and assessment of smaller units of student work (2010:124). Contrary to what the other scholars have explained above, Taras asserts that the term assessment usually refers to a
judgment and it is a process that permeates most of our lives (125). According to her, evaluation is a methodological activity consisting only in “the gathering and combining of performance data with a weighted set of goal scales to yield either comparative or numerical ratings, and in the justification of (a) the data-gathering instruments, (b) the weightings, and (c) the selection of goals” (Ibid.). Taras’s former definitions of assessment and evaluation are in agreement with the definitions from the other scholars in this chapter, but these latter definitions seem to be the opposite from the definitions made by others. Other scholars as seen above and below this paragraph state that it is evaluation, not assessment, which contains “judgment,” while assessment, not evaluation, involves gathering data. This study follows the opinions of the majority of the scholars defining assessment and evaluation.

On the one hand, Burke defines assessment as the process of gathering evidence of student learning to inform instructional decisions (2010:19). His definition also focuses on student learning similar to the definition from Palomba and Banta (1999). Burke further explains that assessment consists of all the tools that teachers use to collect information on student learning and instructional effectiveness. The tools, for instance, include tests, presentations, observations, and class work to assess student learning. On the other hand, he states that evaluation is the procedure for collecting information and making a judgment about it. He provides examples of evaluation products, such as standardized test scores, dropout and graduation rates, and promotion and retention rates used to evaluate the success of a high school. Another difference between assessment and evaluation discussed by Burke is that assessment is an ongoing process occurring daily, while evaluation often happens at the end of an assessment cycle (Ibid.).

In translation, the definitions of assessment and evaluation provided by most TS scholars below focus on translation products, not on translators or translation students. Only Colina (2011) proposes the definitions that are similar to what most scholars in Education
have described above. The word *assessment* is seen by McAlester as an umbrella term, with three specific superordinate components (2000:231). The first component is *translation evaluation*, which reviews the quality of a target text. For example is by giving marks (with either letter or numerical grades) or a binary pass/fail translation. The second is *translation criticism*, which analyzes a target text from the functional standpoint in order to show the appropriateness of a translation or to imply a value judgment. The third is *translation analysis*, which focuses on comparative linguistic analysis involving the explication of the source text – target text relationship, presumably without involving a value judgment (*Ibid.*).

Based on the explanation above, the term *assessment* functions as a broad term in Translation Studies. According to Maier, the terms *assessment* and *evaluation* are sometimes considered synonymous; in fact, many people often use them interchangeably without indicating that they are synonymous (2000:137). Rodríguez also maintains that the terms *assessment* and *evaluation* have often been applied interchangeably without explanation (2007:5). She further states that evaluation is identified with translation quality assessment (see House 1997) and translation criticism (16). In addition, she asserts that there are close boundaries between criticism and evaluation (5). Lambert states that criticism cannot do without evaluation (1991:31), as translation criticism involves the evaluation and/or assessment of the target text based on the source text (Rodríguez, 2007:13). In other words, the basis for translation criticism is the comparison of a target text and its source text (*Ibid.*). Colina (2011), however, claims that there is a clear distinction between assessment and evaluation of translations. She states that “assessment normally refers to a process by which information is collected relative to some known objective or goal,” whereas evaluation is considered to have a subjective component because “when we evaluate, we judge, classifying according to some defined criteria” (2011:43).
The term translation assessment has been chosen for this study instead of translation evaluation based on the definitions in general education (Angelo & Cross, 1993; Straight, 2002; Baehr, 2010; Burke, 2010) and in Translation Studies (Colina, 2011). All those scholars agree that assessment focuses on improving learning, teaching, and future performances, and it provides feedback (formative and process-oriented). All these characteristics of assessment are compatible with the purpose of the study, which is to collect information to determine which of the two assessment models is more effective for translation students in improving their translation skills. Evaluation, however, focuses more on grades (the end results – summative), describes the determination of the quality level, and covers the macro level, such as at the course or university program level. These characteristics of evaluation are not applicable to or beyond the scope of this study.

2.2 The Teacher-Centered Approach to Teaching vs. the Student-Centered Approach

In education generally there has been a shift from a teacher-oriented classroom to a student-oriented classroom. Weimer explains that in educational psychology the focus has clearly changed from the teacher to the learner (2002:13). He indirectly discusses the differences of the teachers’ roles in the teacher-oriented classroom and in the student-oriented classroom. According to him, in the teacher-oriented classroom, teachers play a role as an expert and an authoritarian classroom manager, and they usually improve their teaching skills by developing their presentation skills. In contrast, in the student-oriented classroom, teachers no longer focus on being experts and managers, they lecture less, and they will move around the class more than just stand in front of the class (14). In student-centered learning, there is power sharing between teachers and students, and teachers are less likely to deal with passive, uninterested, disconnected students (31).
Weimer further states that “in learner-centered teaching, instructors guide and facilitate learning” (74). The following are the summarized roles of teachers as described by Weimer. First, teachers can act as gardeners, who prepare the foundation for students to learn. Second, they can act as midwives as they can empower, pull and push, and activate students. Third, they can act as guides, who assist students to improve using some „tools.” Fourth, they can act as coaches, who not only instruct but also design, participate, and make decisions. Fifth, they can also function as maestros, who direct students to produce something (74-76).

According to Schaefer and Zygmont, teacher-centered styles have long been dominant in North America and promote dependent learning (2003:238). In a teacher-oriented classroom, they assert that students become passive recipients of information, and learning is described as taking place if a change in student behavior is evident. However, in a student-centered classroom, the focus is on the needs of the students rather than on the knowledge to be transmitted. Furthermore, they write that “the student-centered approach is characterized by a collaborative relationship between the teacher and the students where all participants are actively engaged in the content and process of learning” (238-239). They also state that there is a foundation of trust between students and teachers facilitating an open dialogue consistent with an exploration of issues in a student-centered learning environment (239).

Lasry et al. state that “traditional classroom settings are teacher-centered” (2013:238). This implies that the modern classroom settings are student/learner-centered. Lasry et al. also repeat what Weimer has discussed before that in a teacher-oriented classroom teachers are placed at the front of the classroom with all students facing them, while in a student-oriented classroom teachers go around the class to encourage students to learn by themselves. Moreover, Lasry et al. explain that a traditional classroom (a teacher-centered classroom) is implicitly based on a „transmission” model of learning where “an expert transmits knowledge to attentive novices” (Ibid.). Lasry et al. do not explicitly discuss what a student-centered
classroom is based on, but in the result of their research they confirm the success of this learner-centered approach in improving active learning by the students because of the interaction between a teacher and students and among students themselves (239).

Similarly, in translation classes, there has been a movement from a teacher-centered approach to a student-centered approach. Tennent (2005:xxiii) states that translation pedagogy has been moving more to a communicative approach or a student-centered approach from a teacher-centered one based on the advances happening in the foreign-language pedagogy. Kiraly (2000) also cites the need to change from a transmissionist classroom (a teacher-centered classroom) to a transformationist classroom (a student-centered classroom) in teaching translation. In general, he sums up by stating that in the transmissionist perspective the teacher should be in control; knowledge is transferred, public, and is considered as a content; learning is molecular and individual; the learner is a student and a client; learning characteristics are shared; and motivation is extrinsic (2000:22). In contrast, with the transformationist perspective, the student should be in control; knowledge is constructed, private, and is considered as a process; learning is holistic and social; the learner is a whole person; every learner is unique; and motivation is intrinsic (Ibid.).

Kiraly states that a traditional translation classroom adopts the teacher-centered approach where it is assumed that there is a transfer of translation knowledge from teacher to student, that the teacher provides an „ideal” translation, and the students are expected to learn from that „ideal” translation (1995:11). Ladmiral in the past (1977) also observed that many translation classrooms still apply the teacher-oriented approach. Similar to Kiraly, he stated that in these classes students” performance is evaluated based on the „ideal” translation provided by the teacher; in other words, students” translation performance is expected to reach the level of their instructor (1977:508). However, Holz-Mänttäri criticizes this teacher-centered approach since it focuses mainly on the grammatical errors of students and it ignores
translation performance (1984:180), although in some teacher-centered classes this might not be the case. Pagano also believes that the teacher-centered practice emphasizes adequacy between the target text and the source text and neglects the acceptability which is necessary to be expressed in the target text (1994:214). Adequacy of a translation refers to the closeness of the text to the source culture, while acceptability of a translation refers to the text’s adherence to target culture (Toury, 1995:56-57).

In contrast to the traditional approach, Holz-Mänttäri believes in the value of student-centered translation teaching because student autonomy from the instructor should be the primary goal of a translation training program. In addition, she states that there is no such thing as an ideal or perfect translation or just one way to produce an adequate translation; instead, it is the responsibility of the translation teachers to assist students to see alternatives, multiple possible translations from any given source text (1984:180). House (1977/1997) also proposes a student-centered approach based on communicative translation in translation classes since this approach will emphasize students’ independence from the instructor, their participation in the selection and production of original texts for translation, and the integration of spoken and written language in the class instruction for translation tasks.

Furthermore, Sainz argues that the central idea in the student-centered approach is that students “learn best when they are involved in developing learning objectives for themselves which are congruent with their current and idealized self-concept” (1994:135). This is more or less similar to providing students’ independence from the teacher. In addition, Sainz mentions two ways in which the student-centered approach contributes to translation teaching. First, she asserts that in this approach students have the right to know the evaluation system used to evaluate their translation and who is assessing their work (138). Second, this implies that this approach requires “a non-aggressive way of giving students feedback on their errors” (Ibid.).
This study ideally should support the student-centered classroom, which might be the wave of the present and the future for translation classes. Asking the respondents (the translation students) their opinions on which of the two assessment models is more effective via a simple survey is one of the efforts to involve the learners/the students in their learning process. However, this study could not thoroughly apply the student-centered approach because, first, I was not the teacher of the class where the projects took place. Second, the projects were conducted remotely as I was in Kent, Ohio, and the respondents were in Jakarta, Indonesia, and this hindered two-way communication between the researcher and the subjects. No discussion could take place like in a normal classroom using the student-centered approach. Nevertheless, in the student-centered classroom, assessment and feedback are ongoing and come from all directions – not just in the instructor’s notes, but also from the work and comments of the other students – and in this study the only aspects that can be observed are the work of the students (the translations and the revisions) and the students’ comments on the translation assessment models.

2.3 Assessment Types

This sub-section discusses types of assessment. In conjunction with the purpose of this observational study, the discussion will focus on pedagogical assessment or assessment for translation teaching. According to Kelly (2005), pedagogical assessment generally takes two major forms: summative and formative assessment. The discussion will start with the differences between summative assessment and formative assessment in education in general. Afterwards, the differences between these two types of assessment in Translation Studies will be explained. There will be a sub-section describing feedback, as this is an integral part of formative assessment. The discussion continues with the definition(s) of certification and
accreditation as a common assessment type in the translation industry. The different definitions of certification and accreditation are also observed from many countries in the world. The last sub-section explains assessment models for writing and two common assessment models for translation.

2.3.1 Summative and Formative Assessment

Several scholars (Bloom et al., 1971; Shepard, 2000; Taras, 2005; Popham, 2008; Burke, 2010) in education cite Scriven (1967) as the first person who introduces the notion of formative in association with assessment. However, his focus is primarily on curricular development rather than on assessing or evaluating the progress of individual students. Scriven explains that summative assessment has always been applied in the curriculum, and he considers this type of curriculum to be unsatisfactory because there is no chance to improve it while it is in progress (1967:43). As a result, he recommends formative evaluation/assessment which involves the collection of appropriate evidence during the process of making or applying a new curriculum instead of waiting until it has been completed. He further maintains that the formative role of evaluation feedback is used internally to improve a program or product, while the summative role of evaluation data is used by outsiders to make final decisions about funding or adopting a program (Ibid.), although not all summative activities are external. Scriven refers to formative evaluation as process (on-going or hermeneutic) research, and its role is to discover strengths and weaknesses in the intermediate versions of a new curriculum (51). From Scriven’s train of thought, we can conclude that summative assessment is conducted at the end of the learning process, whereas formative assessment is conducted during the learning process.
Bloom *et al.* (1971) also discuss the differences between summative assessment and formative assessment. They state that summative assessment is the type of assessment “used at the end of a term, course, or program for purposes of grading, certification, evaluation of progress, or research on the effectiveness of a curriculum, course of study, or educational plan” (1971:117). They consider that the important feature of summative assessment is a judgment made about students, teachers, or a curriculum in connection with the effectiveness of learning or instruction after the learning process has ended, and this judgement has caused anxiety and defensiveness to arise in those who have been assessed. Formative assessment, however, according to Bloom *et al.*, is “the use of systematic evaluation in the process of curriculum construction, teaching, and learning for the purpose of improving any of these three processes.” They also confirm that formative assessment takes place during the formation (construction) phase, and that it should be used to improve the process (*Ibid.*). They consider this type of assessment to be one of the ways to reduce the judgmental aspects of assessment (117-118). Another difference explained by Bloom *et al.* between formative and summative assessment lies in the type of tests used by the two different forms of assessment. They clarify that formative tests are achievement tests of specific units of learning, while summative tests are achievement tests that cover a number of units of learning (137).

Taras argues that all assessment begins with summative assessment, which is a judgment, and that formative assessment is also summative assessment but with feedback used by learners (2005:466). Hence, she emphasizes that feedback\(^1\) is the essential characteristic of formative assessment, as this feedback “indicates the existence of a gap between the actual level of the work being assessed and the required standard,” and formative

---

\(^1\) The detailed explanation of feedback can be found in 2.3.2.
assessment also requires an indication of how some work can be improved to accomplish the required standard (468). Agreeing with Sadler (1989), Taras states that formative assessment also contains judgments just like summative assessment, but it is more about how the judgments can be applied to improve learners’ competence (472).

Taras in another paper (2010) again discusses the differences between summative and formative assessment. She confirms that summative assessment refers to externally accredited exams, but not all summative exams are accredited, while formative assessment refers to any examination or exercise that provides feedback within a classroom (2010:125) or outside a classroom. She further explains that summative assessment is usually in a form of final tests or exams, and it provides information which is called “knowledge of result” (125-126). In contrast, formative assessment “focuses on means, techniques, and procedures to support learning through feedback” (127). Another difference between these two types of assessment according to Taras is that summative assessment produces feedback (which will be applied to future learners), but formative assessment must use feedback (for current students) (Ibid.).

Ainsworth and Viegut (2006) provide more explanation on summative and formative assessment. They refer to formative classroom assessments as pretests or pre-assessments given to students before formal instruction occurs, and those assessments could also be used to improve student progress during instruction and again at its conclusion (2006:23). These formative classroom assessments are also called as assessments for learning by teachers. They further explain that:

Teachers who regularly utilize formative assessments are better able to: (1) determine what standards students already know and to what degree; (2) decide what minor modifications or major changes in instruction they need to make so that all students can succeed in upcoming instruction and on subsequent assessments; (3) create appropriate lessons and activities for groups of learners or individual students; and (4) inform students about their current progress in order to help them set goals for improvement (Ibid.).
On the contrary, summative classroom assessments (given by individual teachers) or common summative assessments (given by teacher teams) are conducted after all instruction and student learning have ended (24). These assessments report the final results of student learning to the teachers, to their students, to students’ parents, and to their administrators, and they support the assignment of letter grades and/or levels of proficiency. Summative assessments are also known as *assessments of learning*. The purpose of these assessments is to measure student understanding of particular units in a textbook, numerous standards during a grading period, or learning objectives for a particular course of study, but the results of these assessments are not used to improve the understanding of the current students. Instead, the results are applied to judge the effectiveness of the teaching practices and to improve instruction for future students (*Ibid.*).

Furthermore, Ainsworth and Viegut confirm that the one broad distinction of these two types of assessment is that if it is formative, the results can be applied to monitor and adjust instruction to improve learning for current students, while if it is summative, the results are for future students (2006:27). In addition, if the assessment is only to measure the final student performance based on what is taught during the quarter, semester, or trimester course of study, then the assessment is summative. If a teacher uses the results from a unit test to inform instruction for the current students during the next unit of study, these test results are being used formatively, although the form of the tests is summative. If a teacher gives students the opportunity to revise and thus improve their performance on a particular assessment during the evaluation process, the assessment can be considered formative. However, Ainsworth and Viegut state that “once the students complete their revision and the final evaluation is determined, the assessment is now summative” (*Ibid.*). In other words, they affirm that when an assessment is the final measurement of student performance, then the assessment must be considered summative (28).
Popham (2008) does not discuss many differences between summative and formative assessment; instead, he focuses more on formative assessment, but his perspective on these two types of assessment is quite interesting. First, Popham repeats the opinion of Scriven (1967) on the distinction between summative and formative assessment. Then, he emphasizes that formative assessment is a way to improve “the caliber of still-underway instructional activities,” whereas summative assessment is a way to determine “the effectiveness of already-completed instructional activities” (2008:4). The rest of the discussion in Popham’s book explains about the characteristics of formative assessment. The important characteristics he describes are (1) formative assessment used to improve student learning is called classroom formative assessment (11), and (2) formative assessment is all about decision making (23). In formative assessment, “teachers decide whether to adjust their current or immediately upcoming instruction in order to improve the effectiveness of that instruction,” and “students decide whether to adjust their learning tactics” (69).

Burke (2010) first repeats the ideas from Scriven (1967), Ainsworth & Viegut (2006), and Popham (2008) on the differences between summative and formative assessment. Then, he explains that formative assessment is assessment for learning because it informs teachers and students, while summative assessment is assessment of learning since it informs the stakeholders (parents, administrators, and public officials) about how well teachers and students have performed (2010:21). In the rest of his book, Burke elaborates summative assessment in greater detail than formative assessment. The following is his explanation of summative assessment.

Summative assessment occurs when teachers evaluate a final product. It usually takes place at the end of a chapter, a unit of study, a benchmark period, a quarter, a course, a semester, or an academic year. Summative assessments report the students’ final results to the students themselves, their parents, and the administration, as well as the school district, the state, and the national government. Summative assessments serve as assessments of learning, because their purpose is to support the assignment of final grades or levels of proficiency related to course outcomes or state standards.
Because summative assessments come at the end of a learning period, they don”t usually include descriptive feedback telling students what they need to do to improve. The only feedback comes in the form of a letter grade, percentage grade, pass/fail grade, or label such as “exceeds standards” or “needs improvement.” (2010:23)

Almost all his description of summative assessment above has been elaborated previously by other scholars, but he explains it more clearly and in greater detail.

Akpan et al. (2012) discuss summative assessment first before explaining formative assessment. They confirm that summative assessment is “the summary of a student”s construction of knowledge,” it is “the evaluation of student performance over time,” and it is conducted “at the end of a unit, course, or grade level to determine the degree of mastery or proficiency according to identified achievement targets” (2012:84). Furthermore, they explain something different from what other scholar have stated about summative assessment, as they affirm that summative assessment provides immediate feedback for teachers and students to assist them to develop a full understanding of the learning process (Ibid.). The previous scholars (Taras, 2005 & 2010; Burke, 2010) do not maintain that summative assessment provides feedback. Akpan et al. further state that formative assessments are procedures conducted as the instruction progresses, and that they are usually short, focused exercises (85). The purpose of formative assessment, according to them, is to “verify that students are keeping pace with the concept and skills presented by instruction”; in other words, it is aimed at skill development at a given point in a specific unit of instruction. They also explain that formative assessment provides periodic feedback and allows teachers to monitor students” immediate progress in the classroom. In addition, formative assessment is used to identify weaknesses in learning before completing a lesson, a module, or a unit (Ibid.).

In Translation Studies, so far there have been only two scholars who discuss the differences between these two types of assessment. According to Kelly, summative
**assessment** is “any assessment activity designed to award grades, permit progression to a higher level, or award a diploma” (2005:161). This kind of assessment is usually conducted at the end of a course period or both in the middle (mid-term exams) and at the end of a semester (final exams). It focuses more on the products or results. Kelly further states that this form of assessment involves grades or a certification (133). Colina claims that “summative evaluation is not meant to include feedback and is therefore less commonly used in translator training programs” (2015:3).

Kelly defines formative assessment as “any assessment activity designed to give learners feedback on their progress and thus enhance learning” (2005:158). This kind of assessment is given throughout the course from the beginning to the end, and it focuses more on the process of learning. According to Kelly, the intention of formative assessment is to consolidate and contribute to learning (134). This type of assessment is considered to be more suitable in a student-oriented classroom because of its role in the learning process (133). Formative assessment provides students with feedback on their learning to assist them to learn more or better (Ibid.). Furthermore, Colina also believes that this type of assessment “is meant to provide feedback during the learning process, so that students can use their mistakes and the teacher’s feedback to improve their learning process” (2015:3).

### 2.3.2 Feedback

As feedback is an important part of formative assessment, its definition needs to be explained here from the points of view of several scholars in different fields of study. The idea of feedback began to exist in economic theory in the UK in the 18th century; however, it was not universally recognized and it did not have a name yet (Mayr, 1989). In a mechanical process, the verb phrase *to feed back* which means *returning to an earlier position* was used
in the US by the 1860s (Jay, 1865; Cole, 1866). In 1909, Karl Ferdinand Braun used the noun word *feedback* as the term referring to *(undesired) coupling between components of an electronic circuit.* The term *feedback* became a distinct word in 1920 because of the action of feeding back of the signal from output to input (Bennett, 1979). According to Ende, the concept of *feedback*, which refers to “information that a system uses to make adjustments in reaching a goal,” was first applied by rocket engineers in the 1940s and has since then been used in many fields (1983:777). In 1950, Norbert Weiner, the father of cybernetics, was one of the first to extend the concept of feedback to the humanities, as he states:

> Feedback is the control of a system by reinserting into the system the results of its performance. If these results are merely used as numerical data for criticism of the system and its regulation, we have the simple feedback of the control engineer. If, however, the information which proceeds backwards from the performance is able to change the general method and pattern of the performance, we have a process which may very well be called learning. (1950:71)

In 1983, Ramaprasad wrote that there was still a lack of a commonly accepted definition of the concept of feedback in management theory, nor was there any universally accepted definition of feedback outside management theory (1983:4). Then he formulated his own definition of feedback, as he states that “feedback is information about the gap between the actual level and the reference level of a system parameter which is used to alter the gap in some way” (*Ibid.*). He explains further that the focus of feedback could be any system parameter (important elements in the feedback system): input, process, or output (5). What he means by *reference level* (in his case) is budgets, pert charts, production, schedules, sales targets, etc. as parameters of an organization (6). The *actual level* refers to the data on the current level of the system parameter obtained by measurement (7).

In general education, feedback is “information that is shared between a teacher and a student, in order to help the student to improve from their current level of achievement, and to achieve their learning goals” ([http://etec.ctlt.ubc.ca/510wiki/Descriptive_Feedback](http://etec.ctlt.ubc.ca/510wiki/Descriptive_Feedback)). By
giving feedback, teachers provide students with advice that guides their learning, helps improve their skills, and assists in closing the gaps in their learning. Besides giving students a grade, it will be more significant to provide them with specific feedback describing their successes and ways to improve their skills, and this kind of feedback is called descriptive feedback. Descriptive feedback can be written, oral, or may even be a question to students while they are working, and it is part of assessment for learning or formative assessment, rather than assessment of learning or summative feedback (Ibid.).

Sadler (1989), Taras (2005, 2010), and Qu & Zhang (2013) discuss feedback in its relation to formative assessment. Sadler states that “feedback requires knowledge of the standard or goal, skills in making multicriterion comparisons, and the development of ways and means for reducing the discrepancy between what is produced and what is aimed for” (1989:142). Sadler also describes feedback in connection with formative assessment by defining that formative feedback is “targeted feedback to improve learning efficiently and expediently” (1989:120). Taras (2005, 2010) emphasizes the importance of feedback in formative assessment because for an assessment to be formative, it requires feedback that “indicates the existence of a gap between the actual level of the work being assessed and the required standard” (2005:468). In addition, Qu & Zhang discuss that a more detailed, continuous feedback will be the outcome of formative assessment, and this feedback will help teachers to decide to use what kind of teaching methods to use based on the different needs of students (2013:337).

In Translation Studies, Washbourne (2014) writes about corrective feedback for a dialogic pedagogy in translator training. He focuses on written feedback and defines it as “one component in an array of potential feedback sources and phases that the translation student or trainee seeks, constructs, and learns from actively” (2014:240). Furthermore, he concludes that feedback cannot be separated from the assignments given to the
students/learners, and that feedback is aimed to improve future tasks (241). Washbourne then discusses the three types of feedback, which he calls as corrective feedback. The first one is direct corrective feedback where the teacher gives students the „right answer(s)” towards the errors they make (242). The second one is indirect corrective feedback where the teacher provides clues that there are errors, but s/he does not give the correct forms for those errors. The third one is metalinguistic corrective feedback where the teacher writes metalinguistic cues towards the errors found, using errors codes and brief grammatical descriptions, depending on the nature of the errors (Ibid.). Washbourne reminds us not to mistake the term corrective in corrective feedback with the meaning of pointing out faults; instead, this term traditionally refers to facilitative (245). The term facilitative means the identification of errors or weaknesses is intended to facilitate self-correction or self-improvement on the part of the student. Error identification is not an end in itself; it is designed to support ongoing improvement on the part of the student. Washbourne then emphasizes that written feedback should be considered “as part of a wider circle of proaction, interaction, and transaction, rather than simply unilateral informing of „what went wrong”” (244). In the end, Washbourne suggested us to systematise feedback via empirical research data (253).

The projects conducted as part of this study have employed two types of assessment that were originally intended as summative assessment – one designed for final exam grading and the other designed for certification testing. Here they are both applied as tools for formative assessment to illuminate the learning process of the participants involved. The two assessment models functioned as formative because the purpose is to support the respondents or the learners in revising their translations, and the feedback applied in the assessment is descriptive indirect corrective feedback. This term is applied here in this dissertation as a combination of descriptive feedback and indirect corrective feedback. The feedback in the projects attempts to improve students” translation skills by describing the strengths of the
students and highlighting on the errors they frequently make. In addition, it also shows ways to correct their translation indirectly; the respondents were given „clues” but not „direct, correct answers.”

According to several scholars in education (Harlan, 2005; Taras, 2005; Kennedy et al.; Lam, 2013) there is still lack of research on how summative assessment might be formatively utilized to improve student learning. In other words, the use of summative assessment as formative assessment or exercises has not been studied to any great extent either in Education or in Translation Studies. The research on modified formative assessment which was originally a summative tool has not been explored much either. The pilot observational study for this dissertation could be one of the pioneers in this research as it uses two originally summative tools to support formative assessment, which in turn supports learning.

2.3.3 Certification and Accreditation

Perhaps the most common summative assessment in the translation industry is conducted in the context of either certification or accreditation, depending on regional usage. Gardner and Palmer define certification and accreditation in terms of conditions in the U.S. According to these authors, in education in general certification is “the process of legal sanction which authorizes the one certified to perform specific services in the public schools of the state” (1982:3). Accreditation, however, is “a private, voluntary process through which an institution is recognized as having met certain criteria and standards,” and “accreditation refers to the peer-based recognition and improvement process for schools, colleges, and departments of education” (4). In other words, certification is for individuals, while accreditation is for organizations and institutions in the U.S.
Stejskal in *ATA Chronicle* (August 2003) discusses the definitions of certification and accreditation in translation. He defines certification as “a voluntary process by which an organization grants recognition to an individual who has met certain predetermined qualification standards,” while accreditation is defined as “a process by which an entity grants public recognition to an organization such as a school, institute, college, program, facility, or company that has met predetermined standards” (2003g:15). Based on these definitions, Stejskal agrees with Gardner and Palmer that certification is for individuals, whereas accreditation is for organizations and institutions. Despite the fact that these definitions could be applied in many countries around the world, there are several countries which use different definitions or even different terms to refer to the same process or activity of (summative) assessment in the translation industry.

According to Stejskal, there are more countries in the world that use the term “certification” to refer to the process to acknowledging the expertise of translators and interpreters than the countries applying any different term. Those countries are the United States, Canada, some of the Arab countries (Kuwait, Syria, Lebanon, Egypt, Morocco), Argentina, Austria, the Netherlands, Finland, Germany, Japan, Mexico, Columbia, Uruguay, Venezuela, Spain, the U.K, Ireland, Ukraine, and the Czech Republic (Stejskal, 2001-2003). However, there are several countries which use the term accreditation to refer to certification as defined here, including Australia, Brazil, and South Africa (Stejskal, 2001a, 2001b, 2001c). Several other countries, Denmark, Sweden, and Norway, apply the term authorization to refer to certification (Stejskal, 2002b, 2002g, 2002h).

In Indonesia, from 1980 to 2010, *Ujian Kualifikasi Penerjemah* (UKP – Translator Qualification Examination) was offered; it was held annually to acknowledge those who passed a test as sworn translators, and the term used for the process is certification, but sometimes it is called accreditation. Unfortunately, since 2011, the UKP has no longer been
held officially by the government via the state university, Universitas Indonesia (the University of Indonesia). Instead, now a translator examination is held several times in a year by Himpunan Penerjemah Indonesia (HPI – the Indonesian Translators Association), but the test validity and formality are still in question. It is unknown why the government has stopped the UKP and when or even whether it will be conducted again.

2.3.4 Assessment Models

Before describing the two common translation assessment models, several writing (essay) assessment models should be discussed first as there are a few similarities between these two types of models. Huot explains that there are three models for the direct assessment of writing quality. The first model is primary trait scoring. Primary trait scoring identifies one or more characteristics (traits) that are relevant to a specific writing task (1990:238). He asserts that “these traits are related to the specific rhetorical situation created by the purpose, audience, and writing assignment.” The second model is analytic scoring which focuses on a few identifiable qualities related to good writing. For example, the quality of a paper is evaluated by how many components of good writing that it possesses. Several analytic scales can be used to assess individual traits from a higher-order concern, such as content, organization, etc., to a lower-order concern, like punctuation (Ibid.).

The third model is holistic scoring that shows “a rater’s general impression of the quality of a piece of writing” (238). In most procedures involving holistic rating, general characteristics represent writing quality for each level of the scale applied. The most popular scales have four to six points (238-239). The characteristics may consist of command of the written convention the language (sentence structure and grammar, for example) (239). Huot also explains that holistic scoring is the fastest, most economical, and therefore most popular
assessment model in writing because a rater just needs 2 to 3 minutes to assess a paper with this model, while it takes a longer time with primary trait scoring and analytic scoring. However, analytic scoring is considered more reliable than holistic scoring (Ibid.).

Another scholar explaining writing assessment models, Kayapinar provides definitions for analytic and holistic assessment models. Analytic assessment is “scoring the levels of the product with individual predetermined criteria and obtaining a total score by the sum of the individual scores” (2014:115). Holistic assessment is “scoring the overall product as a whole, with judging the predetermined component parts separately” (Ibid.). Kayapinar maintains that holistic assessment is more subjective than analytic assessment (114-115). The definitions of analytic scoring from Huot and analytic assessment from Kayapinar are similar as they refer to more granular points or traits. In addition, the definitions of holistic scoring (Huot) and holistic assessment (Kayapinar) are the same. They refer to more general impression or characteristics.

In translation, these two models are more commonly applied than any other models. These two kinds of models are also used worldwide in the translation industry and in translation education, mostly as summative assessment. In the translation industry, the common model used is a checklist of errors, an error analysis/deduction approach, or an analytic evaluation instrument. This model is similar to analytic scoring or analytic assessment described above. Based on the website on Multidimensional Quality Metrics (MQM) Definition developed since 2014, the methods of analytic assessment identify “precise issues within the object being assessed, such as (for a translation) identification of specific mistranslations, spelling errors, etc.” (http://www.qt21.eu/mqm-definition/definition-2015-06-03.html).

This assessment model evaluates a translation product based on the errors it contains. The more errors a translation has, the lower grade it will receive. In other words, the grade is
reduced proportionate to the total errors found in a translation. According to Conde, an analytic evaluation instrument is based on the number of errors and, sometimes, good decisions, whereby these are described, quantified, and subtracted from the total errors (2013:98). This kind of model focuses more on translation products. An example of this model is the ATA Framework for Error Marking, which is used to grade the ATA certification exam\(^2\). This model lists errors mostly proposed by Vinay & Dalbernet (1958/1995), for instance, *mistranslation, addition, omission, literalness*, etc.

This ATA Framework has also been applied to assess students’ translations in universities as formative assessment. Doyle has applied the ATA Framework in a pre- and post-translation assessment (formative assessment) process in his translation classes, although he is not an ATA certification exam grader (2003:23). He provides an example how this assessment model can be used for formative classroom assessment. In addition, Koby and Baer discuss the use of this framework in the translator training classroom as a way to introduce novice translators to the professional standards of error marking (2005:33). They further affirm that this assessment model can be adapted and adjusted into a formative assessment (a process-oriented assessment) (43).

The common model applied in language and translation education provides a kind of *assessment scale* or *a holistic evaluation instrument*, which contains positive and negative descriptions of overall competence. This model is similar to *holistic scoring* or *holistic assessment* described in writing assessment above. Based on the MQM definition, the methods of holistic assessment “focus on overall characteristics of the object being assessed, such as (in the case of translated texts) reader impression, sentiment, clarity, accuracy, style, whether it enables a task to be completed, and so forth” \(^{3}\). Waddington (2001a, 2001b, and 2003) and Turner *et al.*

\(^2\) Further and complete explanation of the ATA Framework is provided in Chapter 4.
al. (2010) apply this assessment scale in their research and call this model a holistic approach because it does not focus on granular errors (see Chapter 3). Based on Conde’s definition, a holistic evaluation instrument classifies “each translation into any of the pre-defined levels within a scale” (2013:98). This model focuses more on the competence of a translator. Waddington and Turner et al. do not clearly state what they actually mean by holistic, but I personally define the holistic approach translation assessment model as assessment which does not focus on granular errors but rather that represents the overall quality of the whole translation, and this definition is based on the research done by Waddington and Turner et al.

In the U.S., an example of an assessment scale is the Interagency Language Roundtable (ILR) Scale (see Appendix F), a model used to assess language and translation competence of those who are applying to work for the U.S. Government. It describes five levels of performance and presents a name for each level and a full description of the competences that the person being assessed must attain to be rated at this level. In some contexts, each level or row in the assessment chart is referred to as a band, such as in the Indonesian Bandscale (the LBI Bandscale) discussed in Chapter 4 on Project Description. As will be evident in the translation assessment development chapter, several authors refer to the individual levels or bands as descriptors, because they are used to describe the competences required at the different levels.

This chapter has discussed the definitions of critical terms which will be used and applied throughout this dissertation. It has explained the differences between assessment and evaluation, and the reasons to choose assessment, instead of evaluation, for this study. Furthermore, the terms teacher-centered classroom and student-centered classroom have been described in relation to one of the types of assessment applied in the study. As I, the researcher, was not really in the classroom while conducting the projects, although I can praise or recommend one classroom approach over the other, I cannot really implement
either. In the sub-section of types of assessment, the differences between summative and formative assessment have been discussed, followed with the explanation of feedback in connection to formative assessment. This study utilizes formative feedback and attempts to develop a model for measuring the effectiveness of the two styles of assessment described here.

The discussion on certification and accreditation aims to assist readers to understand the different use of those two terms in many countries, as several scholars (in Chapter 3) also apply either one or the other of these terms. Similarly, the purpose of the discussion of the assessment models is to clarify that several terms actually have the same meaning or refer to the same thing. It is important to avoid confusion and to understand better what translation scholars all over the world write in terms of certification, accreditation, and the two translation assessment models explained in Chapter 3. Chapter 3 expands on the development of assessment in the field of Translation Studies, starting with Translation Quality Assessment (TQA) by House and then to Skopos theory. Furthermore, the application of Skopos theory in more recent publications in association with assessment is also discussed, continued with the description of theoretical and descriptive theory on translation assessment and empirical research on translation assessment.
CHAPTER 3
TRANSLATION ASSESSMENT DEVELOPMENT

This chapter discusses a range of literature related to translation assessment. It attempts to describe the development of translation assessment and to observe what has been conducted in this field empirically. This translation assessment development firstly can be observed in House’s Translation Quality Assessment (TQA). Then, the most important translation assessment theory is developed from Skopos Theory, profiling Vermeer, Reiss, Nord, and Schöffner, followed by recent publications on Skopos theory related to translation assessment. The discussion on descriptive and/or theoretical publications and several publications involving empirical research on translation assessment is also necessary as part of the rationale to conduct this study.

3.1 Translation Quality Assessment (TQA)

Juliane House was one of the first scholars to specialize in translation quality assessment (TQA). The first edition of her first book, which was based on her Ph.D dissertation, was published in 1977, with subsequent updates in 1981 and 1997. The focus in the book is to provide translation criticism (TQA) with a scientifically-based foundation and to develop TQA as a proper field of study and research in the science of translation. House states that “Evaluating the quality of a translation presupposes a theory of translation. Thus,
different views of translation lead to different concepts of translational quality, and hence
different ways of assessing it” (1977/1997:1).

There are several approaches on evaluating translation discussed in her book. The first set introduces anecdotal, biographical, and neo-hermeneutic approaches, which are “reflected in the century-old, intuitive, and anecdotal judgments of how good or bad somebody finds a translation” and are also known as mentalist views or psycho-social approaches (House, 2014:242). They mostly focus on the process of interpretation in translation (House, 2009:57) and on what happens between the translator and the original text (House, 2014:243). House criticizes these approaches as being too subjective or depending largely on the translators’ subjective interpretation and transfer decisions based on their intuition and experience only, and House maintains that translation is not a „private affair,” as it bears a responsibility to the author, the reader, and the text (1997:3; 2014:243).

The second set comprises response-oriented, behavioural approaches, and the best example is represented by Nida with his Dynamic or Functional Equivalence. They use the reactions of receptors of a translation as the main foundation for evaluating the quality of the translation (House, 2009:44) and mostly do not depend on the source text (House, 2014:244). Given this school of thought, the factor that is considered to contribute to a good translation is an approach that leads to „equivalent response” (2014:243-244). House, however, doubts that this approach can be empirically tested, and she further states “if it cannot be tested, it seems fruitless to postulate the requirement, and the appeal to „equivalence of response” is really of no more value than the philologists” and hermeneuticists” criterion of „capturing the spirit of the original”” (1997:4). In other words, it is still impossible to claim that either of these approaches provides a reliable way to assess translations.

The third set involves text-based approaches, consisting of four individual factors. The first set includes literature-oriented approaches or Descriptive Translation Studies.
These approaches are primarily associated with the work of Toury (1995). They focus on „actual translations,“ which are considered to belong to the important (mostly) literary genre of translation, and on the textual phenomena which are trending in the target culture (House, 2014:245). In this context, the quality of a translation is assessed from the viewpoint of its receptors according to its forms and functions inside the system of the receiving (target) culture and literature (House, 2009:47; 2014:245). The existence of a source text as a basis for the translated text remains very important, but House sees this as a problem because it is hard to determine when a text is a translation and what criteria are used for evaluating it (1997:7).

The second set is made up of post-modernist and deconstructionist approaches. According to House, these theorists attempt to reveal the „unequal power relations” reflected in the translations from and into English (1997:9) and the resulting manipulation in the textual material (2014:246). Here, the source text or the original text is considered to have little importance (House, 2009:47). One of the leading theorists reflecting this trend is Venuti, who attempts to make translations and translators more „visible” and focus on “the „hidden persuaders” in texts whose potentially ulterior, often power-related motives are to be brought into the open” (House, 2014:246). House does not actually criticize these approaches but claims that they place more emphasis on the external factors of the translation action itself, such as psycho-philosophical, socio-political, and ideological factors (Ibid.).

The third set consists of functionalistic and action and reception-theory related approaches, which focus on the purpose of a translation, and the target culture norms are considered the most important factor for assessing the quality of a translation (1997:12). What decides the purpose or function of a translation is the translator or the translation brief (instructions) given to the translator (House, 2014:244). The most notable previous proponents of these approaches are Reiss and Vermeer (1984), and the details of their Skopos
theory are elaborated below. House’s critique towards these approaches is that there is no explicit and operationalized notion of function that is critical to this theory, and that it is not clear how to realize a text’s global „skopos” linguistically (Ibid.). Moreover, House claims that these approaches only emphasize target text factors and ignore the source text, while a translation is bound to both (245).

The fourth set of factors comprises linguistically-oriented approaches. Reiss (1971) is considered as a pioneer in this area, as she assumed that text types (expressive, informative, operative) are the most important invariant for a translation, but House thinks Reiss has failed to provide precise indications on how to conduct an assessment of “whether and how original and translation are equivalent in terms of textual type and otherwise” (2014:247). These approaches focus both on the source text and the target text. House explains that “in these approaches the source text, its linguistic and textual structure and its meaning potential at various levels (including the level of context of situation in a systemic framework), is seen as the most important, indeed constitutive factor in translation” (1997:16 & 2014:247). House’s TQA approach itself can be categorized in this regard (2014:248-252), as she claims that these approaches can provide detailed procedures for analysis and evaluation (247).

In her original book, House presents a detailed non-quantitative, descriptive-explanatory approach to TQA (1977/1997:10-13), and she applies the functional text features developed by Halliday (1978) and Crystal and Davy (1969). House has developed the linguistic, functional-pragmatic model of translation criticism, which is based on equivalence, and this model considers three (semantic, pragmatic, and textual) aspects as important aspects for translation (House, 2014:248). In her revised model (1997), she applies the classic Hallidayan Register concepts of Field, Mode, and Tenor. Field refers to “the topic and content of the text, its subject matter, with differentiations of degrees of generality, specificity, or granularity in lexical items according to rubrics of specialised, general, and
popular” (250). *Tenor* refers to “the nature of the participants, the addresser and the addressee, and the relationship between them in terms of social power and social distance, as well as degree of „emotional charge”.” *Mode* refers to “both the channel – spoken or written and the degree to which potential or real participation is allowed for between writer and reader” (*Ibid.*).

Furthermore, House suggests two types of translation. The first one is *overt translation*, which is a translation that makes the reader aware he/she is reading a translation (1977:189). In this type of translation, the receptors of the translation are not being addressed openly, and an overt translation must be openly a translation, not a „second original”; in addition, its source text “is tied in a specific manner to the source linguaculture” (House, 2014:252). For example, in overt translation the translation text is incorporated inside a new speech event so that it has a new frame (253). The second type of translation is *covert translation*, whereby the translated text functions as an original text in the target text culture (House, 1977:194). This type of translation enjoys the status of an original source text in the target culture, and its source text is not particularly addressed to a specific source culture audience (House, 2014:252-253). In covert translation, the translator tries to recreate an equivalent speech event so that the function of the discourse in the target culture will be the same as the one in the source culture (254). According to House, these two types of translation make very different demands on translation quality assessment (*Ibid.*). Overt translations are more straightforward than covert translations as cultural filtering in these translations can be omitted; overt translations are easier to judge (as translations) than covert translations (254-255).

The explanation on several approaches of translation assessment by House shows that there have been a number of theories on translation assessment. It also reveals a variety of ways to assess translations. House also highlights the strengths and the weaknesses of each
approach. This helps to observe the advantages and disadvantages of the current assessment models, like the ATA Framework and the LBI Bandscale. Of course, there is plenty of room for the assessment theory to develop more or for other ways to do translation assessment to be discovered through further research.

### 3.2 Skopos Theory

In Translation Studies (TS), perhaps the most important current theory related to translation quality and assessment is *skopos* theory, which emphasizes that the translation quality should be measured by the purpose of a translation. *Skopos* theory was developed by Hans Vermeer (1978/2004). The word *skopos* here refers to the aim or purpose of a translation, part of a translational action theory (2004:227). According to Vermeer, “translational action leads to a target text,” and “translation leads to a *translatum*, as a particular variety of target text” (*Ibid.*). In general, Vermeer concludes that “the skopos of a translation is the goal or purpose, defined by the commission and if necessary adjusted by the translator” (236). Vermeer states that a *commission* refers to “one”s own initiative or someone else”s (usually the client) initiative” to initiate a translation job (234), but the translator is the one who determines the purpose for the translation and the scope of the work. The *commission* consists of *translation instructions* or *specifications*. This commission is formulated based on the target culture circumstances, not based on the circumstances of the source culture (235).

Reiss and Vermeer (1984/1991/2013) expanded the skopos theory aiming at a more general translation theory for all texts. The first (1984) and second editions (1991) of their book are written in German with the title *Grundlegung einer allgemeinen Translationstheorie* (*Foundation for a General Theory of Translation*, or as later published in English, *Towards a
General Theory of Translational Action). The third edition published in 2013 is the first edition published in English (translated into English by Nord). The first part of the book was written by Vermeer and contains the theoretical foundations and basic principles of skopos theory as a general theory of translation and interpreting or „translational action.” In the second part, Reiss explains her text-typological approach (first presented in 1971) as a „specific theory” emphasizing cases where the skopos requires „equivalence of functions” between a source text and a target text.

The primary disciple of the skopos theory writing in English, Christiane Nord, focuses on what she calls a functional approach (1984/1991/2005). Her main point is that the intended function or skopos of the target text as determined by the needs of the initiator (commissioner) is what governs the translation process, and it is not the source text, its effect on the ST receivers, nor the function assigned to it by the author (2005:10). This idea is directly based on Vermeer’s Skopostheorie (Reiss & Vermeer 1984:95ff.). In addition, since a translation has to be suitable for a certain purpose, it must fulfil the requirements defined by instructions for translating, i.e., the translation brief (2005:9-10). Nord states that “the translation brief should contain as much information as possible about the situational factors of the prospective TT reception, such as the addressee(s) or possible receivers, time and place of reception, intended medium, etc.” (11). This translation brief is the same as commission introduced by Vermeer (1978/2004).

In addition, Nord claims that translation following functionalism or skopos theory might require a choice of either „faithful” or „free” translation depending on the purpose of the translation. She states that “the law of fidelity in liberty” for translation proclaimed by Benjamin (1972) was not successful in clarifying translation problems (2005:25). She further states that:

The line between fidelity (being faithful) and servility (being too faithful) on the one hand, and liberty (being free) and libertinage (being too free, i.e. adapting or “even” paraphrasing) on the other, is drawn
Thus, according to Nord, it is no longer relevant to assess the quality of a translation based on whether it is too faithful or too free. Ultimately translation quality is a function of the degree to which a translation adheres to the agreed upon brief.

According to Schäffner, functionalism focuses on the purpose of translation to provide a functionally appropriate translation (1997:2). A functionally appropriate translation refers to an (pragmatically) adequate translation, and translation scholars today prefer using this term to using the term a good translation. Schäffner states that “the introduction of the function and/or the purpose of the TT (target text) as the decisive criterion for all translations, and thus, also of TQA (Translation Quality Assessment), is the major contribution of functionalist approaches to translation.” Moreover, she explains that the quality in these approaches “depends on the text user and his/her criteria for assessing how appropriately and efficiently a text fulfils its purpose in a specific situation,” so the quality is relative to the situation rather than being considered „objective” (Ibid.). Schäffner further elaborates that by knowing for whom the translation is being prepared and what the users want to do with the text, the translator can produce a target text that is “appropriately structured and formulated in order to effectively fulfil its intended purpose for its addressees” (3).

In more recent development of skopos theory and functionalism in connection with translation assessment, Secară (2005) discusses different frameworks used in the process of translation evaluation with the focus on error classification schemes applied in the translation industry and in translation teaching institutions. Her paper also describes two recent projects involving translation evaluation. As supported by previous literature (Pym 1992, Sager 1989), Secară believes that there is still no universally accepted evaluation model in the translation
world for the reason that quality is context-dependent (2005:39). She makes a credible case for the idea that there is no single universal translation assessment model because a translation of one language pair might require a different assessment model from the comparable translation of a dissimilar language pair. Moreover, the difference can also be caused by different purposes for assessment. In addition, the assumed intention of the translation itself plays a role. If the intention of a text is to inform, one set of criteria applies. If it is to persuade, another set comes into play.

Colina (2008, 2009) proposes a componential-functionalist approach to translation quality evaluation. The approach is referred to as componential since it evaluates quality components separately, and as functionalist “because evaluation is carried out relative to the function specified for the translated text” (2008:97). Colina made this evaluation proposal to aid US health providers in using translated materials, and these materials are assessed with this approach (105). As the object of assessment is a product (translated materials), the assessment is considered summative (Ibid.). The approach applies descriptive statements (descriptors) rather than a point deduction system since the raters must choose the descriptors which best describe the text under assessment in several areas (106). This approach, which she refers to as a tool, has been tested, and the results show that it exhibits good inter-rater reliability (2009:235). Based on her research, Colina also concludes that both professional translators and language teachers are equally qualified to assess translation quality using this tool (255). In brief, Colina states that the tool assesses the overall quality of a translation by evaluating diverse components of a translated text separately (Colina, 2015).

Polliastri and Paulina focus on translation competence and how it can be measured. According to Pym, translation competence is “the knowledge, skills, and attitudes necessary to become a translator” (2011:78). Polliastri and Paulina state that “translation quality is a key issue in the inter-linguistic/intercultural communication market as well as in
translator education environment” (2009:239), and they maintain that quality is complex, and only some of its components can be measured (240). It will be impossible to assess translation quality from all factors involved. They further suggest some categories for translation quality criteria, such as accuracy (content and meaning), creativity (appropriate to text types – see Reiss 1981/2000b – and client needs), and skopos (translation goals/intentions) (246).

From all those criteria, only one of them can be applied in this observational study, and that criterion is accuracy, and it will be addressed by comparing errors in translation. To assess creativity and skopos criteria, the respondents must translate different types of texts. For example, the respondents are asked to translate a journalistic text and a scientific text, and if they can show in their translations the differences between a journalistic text and a scientific text, it means they are creative, able to translate according to the text type, and to follow the purpose (skopos) of the translation task. However, in this study there is only one type of text to be translated, a non-technical text (see Chapter 5), and there are already three tasks that the respondent must do in the research. Furthermore, testing other criteria would complicate the study and require additional research. The translation and revision tasks given to the respondents in this study are provided with translation brief (see Appendix D), which is a direct influence from skopos theory.

3.3 Theoretical and Descriptive Theory on Translation Assessment

Several publications on translation assessment/evaluation contain theoretical and/or descriptive discussions. Kasandrinou (2010) provides a descriptive study about translation evaluation and looks into how evaluation is usually conducted and its significance for most translator training institutions. She reports that “although there are inherent differences
between industry/commerce and translation education, the principles of QA (Quality Assurance) can prove extremely useful when transferred in the translation training environment” (2010:194). Her statement is quite sensible, as one of the aims of translation education is to prepare novice translators to become professionals in the translation industry. Furthermore, she asserts that “evaluation should be seen as a tool to gather information which will be then used to improve practice” (195).

Orlando explains and justifies the use of two types of evaluation tools designed and implemented in different contexts and at different stages of the training of future professional translators, at Monash University, Melbourne, Australia (2011:293). The two different evaluation tools applied in this university are the (1) NAATI (National Accreditation Authority for Translators and Interpreters) assessment criteria (see Appendix H) to evaluate translations in a product-oriented way, and (2) a translator’s diary (including translation drafts), which is a more formative, student-centered approach (295). Orlando claims that the purpose of these different tools is to make translation teaching and learning become more effective, as he states, “In a program aiming at training future professionals along theoretical and practical lines, students and instructors must all understand that there is not just one way to evaluate translations” (296-297). Personally, I think it is understandable that there are several ways to assess translations depending on the purpose of the assessment and on the language pair of the translation. For example, the assessment of a legal text should be different from the assessment of a literary text since those two texts have completely different characteristics. Furthermore, English-Indonesian translation assessment may require a different assessment strategy from Indonesian-English translation assessment.

Marais (2013) objects to the idea of a unified theory of assessment in translator education following Toury”s (1995) theory of translation, which is based on the lack of consensus on a definition of translation itself. He argues that “the very natures of translation
and assessment make it impossible to conceptualize one universally valid method for the assessment of translation” (2013:14). Parallel with what Orlando has described above and with the fact that there is no universal theory of assessment, Marais also claims that there is no such thing as a universal way to assess translations, and that it is not viable to create one ideal instrument for translation assessment (Ibid.). Moreover, although a significant amount of work has been conducted on translation assessment over several decades, Marais asserts that there are still very little agreement in the field of translation assessment, a lack of reliability in assessment instruments, and a lack of rigor in assessment (13). His argument is that “no matter how clearly we define assessment, evaluation, or certification, the particular instrument used in whatever we call the process will be theoretically biased” (Ibid.). Like House, he asserts that every translation assessment is explicitly or implicitly based on a certain translation theory (14). Hence, we should not apply just one type of translation assessment, and we should be open-minded towards different types of assessment. In fact, a combination of assessment types might lead to more effectiveness than using merely one translation assessment type.

### 3.4 Empirical Research on Translation Assessment

Waddington (2001a, 2001b, and 2003) reports on the results of empirical research in translation assessment conducted while writing his dissertation. The focus in all his papers is on summative assessment. In his research, he examined student translation results assessed using several approaches or four assessment models. His experiment involves two different approaches to error analysis, consisting generally of the use of error checklists. Another assessment model used in Waddington’s research he calls a holistic approach (similar to an assessment scale) which uses a unitary scale treating translation competence as a whole, does
not involve an error checklist, and has descriptions of positive and negative criteria of translation results written in a clear and simple language (2001a:21). Furthermore, a fourth model is comprised of the combination of error analysis and his holistic approach. Waddington’s overall results reveal that all methods turned out to be equally valid or there is no doubt of the validity of all the methods applied, either based on error analysis or on a holistic approach (2001a:23, 2001b:311), yet he concludes that the error analysis-based methods are more reliable than the holistic one (2001a:32) because they produce marginally greater accuracy than a holistic approach, but the combination of both shows even greater accuracy (33-34).

In a later paper (2003), he focuses on three factors. The first examines an approach involving the correction of student translations which considers the effects of mistakes on overall quality (2003:410). The second focus is on the results gained from this method compared with the results of other methods. The third factor uses a positive approach towards the mistakes students make, as the mistakes can help measure source text difficulty and reveal what the role of the specific mistakes is in the learning process (Ibid.). This third focus is particularly interesting, as Séguinot\(^3\) states that errors can provide us information on translation quality and translation process (1990:68). From the research results, Waddington concludes that mistakes made by students should be considered as potential information on translation student progress and on the learning process of doing translation (2003:424).

Conde in his dissertation (2009a) compared evaluations carried out by professional translators, translation teachers, translation students, and potential addressees. He confirms Waddington’s results (2001a & 2001b) that holistic evaluators reached the same conclusions as analytic evaluators, and he asserts that holistic models were as effective as error analysis models for measuring translation quality. Moreover, Conde’s research implies that quality

\(^3\)More discussion on Séguinot can be found in Chapter 6.
judgment does not always depend on how detailed the evaluation is (2009a:331). Aside from his dissertation, Conde has published several papers, all of which are based on empirical research. The first paper he published in 2009 analyzes translation evaluation in the teaching environment. Overall this paper supports continuous assessment, blind evaluation\(^4\), and holistic systems, and the assessment of learning\(^5\) (2009b:231).

Next, Conde in another article introduces the relationship between translation evaluation and certain features on the surface of texts (2011:69). He conducted research with 88 subjects, and those subjects or respondents were asked to assess 48 texts. They were not asked to do translations of the texts, but they were requested to do assessment on 48 different translated texts. It is not known in this paper who translated the texts. The features or aspects Conde explored are emphasized words, actions on emphasized words, text length, resemblance to original format, lexical density, and readability. Even though Conde does not explicitly define what emphasized words are, it can be implied that *emphasized words* are the words which are considered important in the texts translated. Conde attempts to discover whether those aspects mentioned have any connection with the increase in the attention levels of evaluators (*Ibid.*). His results, which he claims are still preliminary, reveal that lexical density (measured by the type/token ratio) and readability (determined using indexes of words and sentence length) do not seem to affect the evaluators’ average quality judgment and are also not directly related to quality (83). Furthermore, “the clearest effect is that related to productivity: the more complete the translations, the higher the judgments issued by the evaluators” (*Ibid.*). In other words, when a student manages to complete a translation, the evaluator is inclined to grant a more positive evaluation. When the translation is not completed, it will be subject to more negative evaluation.

\(^4\) Blind evaluation is the evaluation where the results are concealed from the tester or the testee or both to avoid bias.

\(^5\) Assessment of learning is the assessment that becomes public and results in statements or symbols about how well students are learning.
Conde also focuses on assessment to seek the relationship between quality and quantity and their effect on empirical data (2012a:69). He states that there are many analytic systems (error analysis/deduction systems) used every day to assess translation “based on subjective experience backed up by very little empirical evidence” (67-68). In this article Conde puts forward some criticisms of analytic systems, one of which is that “these approaches emphasize the role of contrastive grammar and linguistics at the expense of factors such as communicative situation and success” (68). Furthermore, Conde claims that many teachers have been using analytic systems in an “excessively individualistic way,” while there is no universal method for these approaches (Ibid.). However, he admits that analytic evaluation methods are considered to be more informative and thorough than holistic ones (71). Overall this paper analyzes three areas where quantity and quality issues are related to the field of translation evaluation: (1) the potential relationship between the detailed level of evaluators and the quality judgment they provided, (2) the relationship between nature of phenomena and saliency\(^6\), and (3) the impact of the most salient phenomena of the texts or quality judgment (78-79).

In a 2013 article, Conde compares translation errors and language errors, and investigates whether both have impacts on the overall quality judgment attached to a translation (2013:97). He defines language errors as errors which can be detected in just reading the target text and that contain mistakes in vocabulary, syntax, grammar, punctuation, coherence, style, etc. (98). Translation errors can be found by reading the source text, and they are usually called errors of meaning (Ibid.). His research results show that “language errors are much more common in translation evaluation; however, translation errors contribute more to the variability of the marks issued” (97). In professional environments,

\(^6\) The **salience** (also called **saliency**) of an item – be it an object, a person, a pixel, etc. – is the state or quality by which it stands out relative to its neighbors. Salience (language) = the property of being noticeable or important.
translation errors play a more important role in assessing translation quality, while in teaching environments, both translation and language errors have an equally significant role in quality judgment (107). The level in which Conde was conducting the research is the undergraduate level. The work of the undergraduate level often reflects the fact that most of the students are still at the formative acquisition of the second language.

Turner et al. (2010) examine two assessment methodologies used for translating and interpreting accreditation tests: error analysis/deduction and descriptors. Based on their research, they discover that there are 3 models used in accreditation systems: (1) error analysis/deduction, (2) descriptors (descriptions – see Chapter 2: criterion referencing, the use of scales to describe performance in tests and providing positive and negative criteria for translation results) with no error analysis/deduction, and (3) a combination of descriptors and error analysis/deduction (2010:12). The first model is usually applied by some translator’s associations, such as NAATI in Australia and ATA (American Translators Association), while the use of the second model (descriptors) has been pioneered by organizations that assess general language competence, such as in IELTS (International English Language Testing System) (15). This model (see Appendix G) is considered to be a holistic system because it describes the way test candidates use language beyond the level of a particular word or grammatical category which can be assessed as correct or incorrect (Ibid.). The result of this paper implies that both error analysis and a descriptor system may be as reliable and accountable (22), even though at the beginning of their paper the authors tend to be in favor of the descriptor system. Their data indicate that error analysis or the descriptors show about the same level of accuracy.

In brief, this chapter has explored translation assessment development written by many translation scholars around the world. Starting with TQA (Translation Quality

---

7 See Chapter 2 in the subchapter of Certification and Accreditation. In Australia, accreditation has the same meaning as certification.
Assessment) pioneered by House and continued with Skopos theory as the foundation of translation assessment theory, we can see that the discussion has developed from theoretical and descriptive considerations to empirical ones. From all the publications involving empirical research, the studies focus only on summative translation assessment. There has not been any study which is conducted on formative translation assessment or formative exercises, or on the reuse of approaches or instruments designed for summative assessment but applied as formative exercises. This is where the present observational study fills in the gap. It observes how effective two originally summative assessment tools are in assessing formative exercises and attempts to discover which one of the two models is more effective by comparing the errors that the respondents made.

Chapter 4 on Project Description will describe this observational study, explaining its design, structure, and location. Then, what LBI Translation classes are will be explained. Next, the two assessment models, the LBI Bandscale and the ATA Framework, are discussed. Further discussion will be about the differences between the LBI Bandscale and the ATA Framework, and about the similarities and differences between this research and other research conducted by others described in this chapter.
CHAPTER 4
PROJECT DESCRIPTION

This chapter will discuss this observational study in detail, starting with the explanation of why the study is called *pilot observational study*. Then it will continue with the reasons why the research was carried out in LBI Translation classes in Jakarta, Indonesia, and why it was conducted remotely. Furthermore, there will be elaboration of what LBI Translation classes are with some history on the establishment of LBI. In addition, the two assessment models, the LBI Bandscale and the ATA Framework, will be discussed separately, comprehensively, and respectively, accompanied by a discussion of the differences between the LBI Bandscale and the ATA Framework. Finally, the similarities and differences between this study and other studies conducted by others described in Chapter 3 will be explained.

4.1 The Pilot Observational Study

The research conducted for this study has been characterized as a pilot observational study for two reasons. The first is that this study only observes the effectiveness or the impact of two assessment models in improving learners’ translations. The effectiveness is seen from the improvement the respondents make in their revisions or the second versions of their translations. The improvement here is reflected in the error mitigation or the error frequency they display in their revisions. If the feedback associated with using a translation assessment
model can lead to fewer errors in a revision, it means this model is effective to assist learners in improving their translations. The study also attempts to observe which one of the two assessment models is more effective. The two models used in the projects are different in many ways; one is an error analysis, and another is an assessment scale (see Chapter 2 in the subchapter of assessment models). The detailed explanation of the models can be seen below. In addition, the study attempts to correlate the learners’ opinions on translation assessment via a simple survey with their translation results.

The second reason why this research is called a pilot observational study is because it is hoped that this study will lead to further research which will involve more respondents, more time, and more variables. The present study collected the data over two years (five terms of LBI classes), but the number of subjects was not and could not be large enough to generate scientific research results due to the limited number of available students and the time constraints involved in this dissertation writing process. The results, however, may provide some indications that highlight the effectiveness of translation assessment models and some indications as to which of the two models is more effective. These indications will be the beginning and the foundation of further, more reliable scientific results in future research. According to Muñoz and Conde (2007), although some results may be just speculations or inclinations due to a small number of respondents, they can still be fruitful for translation assessment development.

The study started with pilot projects during two terms of LBI classes to test the feasibility and the validity of the variables involved in the study. The formal or second projects were conducted during three terms after several changes were made based on the outcomes of the pilot projects. All projects basically involved translation students or learners in LBI translation classes. Those respondents were asked to work on a translation at home for a few hours, and the translation was submitted to me via email. After more or less two weeks,
the respondents’ translations were returned with assessment and feedback using two assessment models, the LBI Bandscale and the ATA Framework. Then, the respondents were asked to do revisions on their translations based on the assessment and feedback given. The errors in the revisions were compared with the errors made in the first translation. During the formal projects, an independent judge was also engaged in assessing the revisions in order to avoid or demonstrate potential biases. The last step of both projects asked the respondents to fill in a simple survey to discover their perspectives on translation assessment. Project details are discussed in Chapter 5, Methodology.

The projects were conducted using subjects who were enrolled in LBI translation classes in Jakarta, Indonesia. The first reason for enlisting these particular subjects is that LBI is the institution where I work so it was easy for me to obtain permission from the director of LBI to conduct research in its classes. Second, translation classes are not very common in Indonesia, and LBI is one of very few language institutions in Indonesia that conduct translator training. Third, LBI students tend to be people who are truly eager to learn about translation and improve their translation skills. The last reason is the study involves the Indonesian-English language pair, which is advantageous because it would be impossible to conduct such research in the U.S. working with advisors and readers who do not read Indonesian, and I am unable to work with any of the languages common to Kent State students (Arabic, French, German, Japanese, Russian, and Spanish). As a result, the study had to be carried out remotely, as I was in Kent, and my respondents were in Jakarta. The detailed information of the respondents is discussed in Chapter 5.
4.2 LBI Translation Classes

LBI stands for *Lembaga Bahasa International* (International Language Institution). It is one of the programs under the Faculty of Humanities (FIB – before 2002 the name was FSUI, the Faculty of Letters) of the University of Indonesia (UI), one of the best universities in Indonesia. Under LBI, there are three centers: *Bahasa Indonesia untuk Penutur Asing* (BIPA – Indonesian Language Teaching for Foreigners), *Program Pelayanan Bahasa* (PPB – Language Service Program, which teaches foreign languages to Indonesians), and *Pusat Penerjemahan* (PP – Translation Center). In 1991, FSUI established *Program Pendidikan Penerjemahan dan Juru Bahasa* (PPPJ – Translation and Interpreting Education Program), and in 1994, PPPJ became PP. In 2008, FIB UI decided to unite all its business ventures under the same roof, so that was when LBI was established.

PP LBI UI offers several services: (1) translation and interpreting services, (2) certification for translators from 1991 to 2010, and (3) translation classes. The translation classes consist of legal translation classes and general translation classes. Legal translation classes have three levels: a basic legal translation class, an intermediate legal translation class, and an upper-intermediate legal translation class. The general classes consist of an English-Indonesian translation class and an Indonesian-English translation class. The general classes are usually taught by a team of teaching staff, with 5 to 7 teachers. Each of them teaches a different type of translation, such as literary translation, journalistic translation, academic text translation, and others. In one year, there are three terms of the LBI translation classes. Term 1 is from January to April (Spring), Term 2 is from May to August (Summer), and Term 3 is from September to December (Fall).

The translation classes at LBI are called *translator training*, instead of *translation education*, because these classes focus primarily on practical topics and provide very little
theory of translation. When translation theory is taught, it is called „tips“ for translation. In addition, these classes are considered informal and for the public, and do not produce any university degree. They are not part of a programmatically planned curriculum or of any sequenced courses. The Indonesian-English translation class is considered as a higher level class which requires more skills from the students than the English-Indonesian translation class, but it does not mean the participants passing the English-Indonesian translation class are required to continue to the Indonesian-English translation class. A participant can enroll directly in the Indonesian-English translation class without attending the English-Indonesian translation class first if his or her language skills meet the requirements. However, if a participant passes the English-Indonesian translation class, he or she can automatically attend the Indonesian-English translation class (if he or she wishes to join the class).

The students of LBI translation classes are very much motivated to learn translation. The first reason is that these classes are not mandatory for anyone and are not related to any program, so people enrolled to these classes realize that their sole purpose is to improve their translation skills. Second, the fee for these classes is one of the most expensive for public courses or training in relation to languages in Indonesia. Not many people living in Jakarta, the capital city of Indonesia, can afford them. It takes significant interest for the people to join these classes because of the time, energy, and money that are required. Thus, this also explains why there are not many students enrolled to these classes in every term (maximum only 15 people but this is very rare). Therefore, the students of the LBI translation classes are excellent respondents for this study since each project of the study consists of several tasks which might be completed only by those who are interested in translation and highly motivated to improve their skills. These respondents are perfect as subjects to seek out the effectiveness of the two assessment models, and based on the results (see Chapter 6) the
respondents show enthusiasm for the study and for the two assessment tools. The following two subchapters will discuss the two assessment models used in the projects.

4.3 The LBI Bandscale

As noted above, the projects that make up the study utilize two assessment tools to demonstrate their effectiveness in assisting translation students to revise their translations (formative exercises). According to Shepard, assessment will be more helpful for students to improve their skills if it is part of the teaching and learning process or formative exercises than just as “the end-point of instruction” or summative assessment (2000:10). The first instrument is the LBI Bandscale, which was formulated in 2008 by the coordinator of translation courses and the PP LBI Manager at that time, and it is still used today. The bandscale is used for the summative assessment of final student translations to evaluate their performance, but the students have never seen the assessment of their final work, nor have they ever prepared revisions of their translations based on this assessment.

The LBI Bandscale consists of four levels of measurement, using the letters A, B, C, and D to reflect student achievement. It is very simple and similar to the descriptors described by Turner et al. (2010) or “a holistic approach” used by Waddington (2001a, 2001b, 2003). As LBI is basically a language institution, the assessment model chosen by the PP LBI UI is the one closer to the model used to assess language competence. The following is the explanation from Turner et al. about the origin of “descriptors” before it was applied to also assess translation.

The use of descriptors in the language field has been pioneered by organizations that assess general language competence. The introduction of holistic systems such as the IELTS has resulted in greater sophistication in language assessment. Unlike previous error analysis or intuitive and non-standardised means of assessing language, IELTS and similar descriptors describe the way test candidates use language beyond the level of a particular word or grammatical category which
can be assessed as right or wrong. Importantly, they describe what a candidate can do as well as what they cannot do (Turner et al., 2010:15).

Because translation in LBI UI particularly and in Indonesia in general is still considered as part of language skills and competence, it is understandable that this institution has chosen an assessment model that is similar to the one applied to assess language. Furthermore, we can see from the quotation above that this assessment model focuses more on the competence of the translators than on the translation products.

The comments for each band in the LBI Bandscale describe not only the errors the candidates make, but also the positive aspects of their translations. Those comments mostly focus on comprehension of the source text and writing in the target text. The comments also serve as fixed feedback to the translation results of the participants, so that the raters do not have to write individual feedback anymore. For instance, when a translation is very good, and a rater gives A, the student producing this translation just needs to refer to the comment in the column next to A grade as the feedback. However, in order to apply this model to assess formative exercises, it has to be modified. Before deciding what scale the students’ translations belong to, all errors found will be marked in their translation but without the explanation of the types of errors since this model does not provide the system to do so. As a consequence, students in the project will see where they have made errors and they will receive an overall evaluation, but no detailed error-specific feedback. The details of this bandscale can be seen in Appendix A.

4.4 The ATA Framework

The second tool is the most current version of the ATA certification assessment framework, which is complex and provides a flexible instrument comprising a detailed,
metric-driven error checklist. It consists of three components: a weighted matrix for error checking, a chart listing the error names (labels) plus descriptions of the individual errors, and a flowchart providing guidance on weighting the errors. According to Doyle (2003) and Koby and Baer (2005), although the ATA Framework was developed for certification, it can also be used as the method of evaluation or grading accompanied by feedback in a translator training. Doyle states that “the framework provides a ready-made, standardized, time-tested, and professionally recognized model for conducting theory-based, systematic, coherent, and consistent evaluations of student translations” (2003:21).

Furthermore, he maintains that any instructor can adopt and adapt this framework in his/her translation classes, even if he/she is not an ATA certification exam grader (2003:23). Even though the framework is used to assess the end product in a translation certification exam, Koby and Baer claim it is possible to adapt and adjust the categories of the ATA error marking scale to transform a product-oriented scale into a more process-oriented one (2005:43), which is one of the characteristics of formative assessment. The ATA Framework consists of two main parts: (1) translation/strategic/transfer errors and (2) mechanical errors. The former set focuses on the negative impacts on the understanding/use of the target text, and the latter focuses on the negative impacts on overall quality of the target text. The second component of the ATA Framework relegates pure language and grammatical errors to this part of the framework, which for professional translators is considered less significant than the real translation errors addressed in the first part of the framework. The details of this ATA Framework can be seen in Appendix B.

The categories in the ATA error marking scale reflect the theory of Vinay and Dalbernet (V & D) (1958/1995). V & D’s principles describe interlingual transfer and have become widely accepted. They were the first translation theorists to introduce translation procedures and strategies as well as translation errors. The list of “procedures” produced by
Vinay and Darbelnet consists of seven categories: *emprunt, calque*, and *traduction littérale* (defined as “direct” procedures), and *transposition, modulation, équivalence*, and *adaptation* (defined as “oblique” procedures) (1958:31-41). Besides these “procedures,” Vinay and Darbelnet listed a few more procedures in the glossary preceding their book (Muñoz, 2000:131). According to Delisle, Vinay and Darbelnet’s translation procedures are more suitable as “labels” which are attached to results, and they are not procedures that can assist translators in doing translation (1988:72-72). Therefore, these “labels” can be applied well in a translation assessment model based on error deduction because the focus of this model is the translation results/products.

4.5 Further Differences between the LBI Bandscale and the ATA Framework

The ATA Framework displays detailed error analysis, while the LBI Bandscale attempts to incorporate both positive and negative criteria of a translation product. Assessment with the ATA Framework will be accompanied by indirect feedback tailored to the performance of each student in order to make it suitable to assess formative exercises. As described in detail in Chapter 2, Washbourne (2014) writes about three types of feedback for translation pedagogy. The first one is *direct corrective feedback* where the teacher gives the students the correct form (2014:242). The second one is *indirect corrective feedback* where the teacher does not provide the correct form, but he/she indicates there is an error. The third one is *metalinguistic corrective feedback* where the teacher gives “some kind of metalinguistic cue as to the nature of the error,” such as providing the use of error codes or brief grammatical descriptions (*Ibid.*).

However, assessment with the LBI Bandscale does not support feedback for each student. The feedback only refers to the explanation of the scales assigned. Consequently, it
is more work for an assessor to apply the ATA Framework than to apply the LBI Bandscale as formative translation assessment. The use of feedback, either tailored or fixed, is necessary to transform the function of these two assessment models from summative assessment to formative assessment. Washbourne states that “feedback in translation conceived as a feedback loop engages the student with standards, and communicates in a way that helps him or her to monitor and improve performance” (2014:243). Thus, it is obvious that feedback can assist students, in this case translation students, understand how to improve their skills, and they can show the improvement through their revision or the second draft of their translation.

Based on Orlando (2011), introducing different assessment models to novice translators studying at the PP LBI helps make them aware that there are several ways to assess a translation. For Indonesian novice translators, the LBI Bandscale is a form of a local assessment model, while the ATA Framework is a foreign assessment model. As explained before, the LBI Bandscale follows the language assessment model like IELTS because translation is still considered part of language skills and competence. Neither the ATA Framework nor any other assessment models using error deduction are common in Indonesia. One purpose of this observational study would be to introduce this “new” type of translation assessment to Indonesian novice translators in LBI in particular and professional translators in Indonesia in general by showing its effectiveness compared with a translation assessment model using “descriptors” or “scales.”

As also described in Chapter 3, Conde states that there is very little empirical evidence to support the use of many analytic systems (error analysis/deduction systems) to assess translations; hence, the use of an analytic system is based only on subjective experience (2012:67-68). This concern may apply in the use of the ATA Framework as a formative feedback tool because it was not originally designed for this purpose. As a
summative feedback tool, this model has gone through some research to prove its effectiveness. However, holistic approaches, such as the LBI Bandscale, have never been proven to be effective through empirical research either, and the LBI Bandscale has never been used to assess formative translation exercises. Therefore, this observational study can become a pilot study for further empirical research to reveal the effectiveness or ineffectiveness of both models as formative assessment.

4.6 Similarities and Differences between This Study and Other Research in Translation Assessment

There are several similarities between this study and the research carried out by Waddington and Turner et al. First, all this research focuses on evaluating student translations. Waddington’s research applied translation assessment to the correction of a second year translation exam into a foreign language (Spanish-English) done by 64 university students. Turner et al. conducted the research on RMIT students with several language pairs. This study involved the students of the LBI translation classes also translating into a foreign language (Indonesian-English). The second similarity is that all of the projects compare an error analysis approach with a holistic approach (an assessment scale). Waddington, Turner et al., and I compared these two types of assessment to discover their effectiveness.

However, there are several obvious differences. First, both Waddington and Turner et al. focus on summative tests, while this study concentrates on formative assignments. Waddington and Turner et al. requested their respondents to work on a translation only, whereas in this study the respondents were asked to first do a translation and then to revise their first draft based on the assessment and feedback given to them. Another difference between Waddington and Turner et al.’s research and this study is that the study attempts to
establish a correlation between student perspectives from their responses in a simple online survey and their translation results checked by me (Rater 1) and an independent judge (Rater 2) in addition to comparing two assessment models. However, Waddington and Turner et al. emphasize only assessment effectiveness by comparing different methods.

Another difference is the variety of assessment models used in the study. Waddington compares two error analysis approaches, one holistic approach, and one combination of both approaches. Turner et al. compare assessment outcomes at RMIT (The Royal Melbourne Institute of Technology) between the use of descriptors (used in Diploma Public Service Interpreting at the Institute of Linguistics in the UK and taken from http://www.iol.org.uk/qualifications/DPSI?DPSIHAndbook.pdf) and the current NAATI error analysis/deduction system. This study compares only one error analysis approach (the ATA Framework) and one “holistic” approach or an assessment scale (the LBI Bandscale). The last difference is, of course, the different language pairs. Waddington worked on Spanish-English translations, and Turner et al. worked on English-Chinese translations, English-Japanese translations, Arabic-English translations, and Italian-English translations, while this study focuses on Indonesian-English translations.

In these research projects, an assessment method as a certification tool (the ATA Framework), which is related more to the translation industry than to translation education, is applied in translation education (see Kasandrinou 2010). As described before, these projects attempt to compare the effectiveness of the two assessment models, and the effectiveness here is based on student improvement in response to assessment and feedback. Student improvement means the revisions will have fewer mistakes/errors than the previous work or the first draft of their translation. In addition, as these projects also attempt to establish a correlation between student perspectives on translation assessment tools and their translation results, a simple survey was prepared for the respondents. According to Terran et al., the use
of such a survey or a questionnaire in translator training is still rare (2012:463), and the research involving student perspectives is even scarcer, although translation students are also major stakeholder in translator training (Li, 2002: 513).

Besides the description above, Brown et al. argue that whether a class is a student-centered or a teacher-centered, “students are not passive, tabula rasa recipient-responders to assessment; they have their own understandings, beliefs, attitudes, opinions, and perspectives of and about what they are participating in” (2009:3). This means that whatever the class approach is, it is important to acknowledge students’ opinions, especially on assessment, in order to reveal whether assessment is congruent with their expectations and needs. In this research, it is significant to discover the effectiveness of any one type of translation assessment in helping students improve their translation work and to reveal their perspectives on translation assessment.

This chapter has elaborated this pilot observational study thoroughly from the reasons why it is called observational study, what the study is, how it is structured, the reasons it was conducted in LBI translation classes and remotely to the explanation of what the LBI translation classes are and the two assessment tools used in the study. In addition, the similarities and the differences between the LBI Bandscale and the ATA Framework and between this research and other research by several scholars described in Chapter 3 have been discussed to provide deeper understanding of this study. The purpose of this detailed description has been to clarify the explanation of the study methodology that follows in the next chapter, which outlines the procedure, participants, materials, and survey involved in the methodology used for the study projects.
CHAPTER 5
METHODOLOGY

This chapter elaborates the methodology of the study. It provides a thorough explanation on how the projects were conducted and a detailed description of all the components involved. The explanation of the pilot projects will be separated from the explanation of the formal or second projects. Each project discussion includes the procedures, the participants, the materials, and the survey. The changes in the procedures, the materials, and the survey from the pilot projects to the formal projects will be elaborated as well.

5.1 The Pilot Projects

The pilot projects were conducted in order to discover whether the design of the study was feasible, valid, and reliable. In any empirical research, carrying out a pilot study is highly desirable before proceeding to the formal research so that there will be a chance to improve the design of the study to obtain more valid and reliable data (Heiman, 2001:89). This subchapter first explains the procedures or how the study was conducted. Then, it discusses the participants in these pilot projects, which were carried out in the Fall 2013 and the Spring 2014. Finally, it elaborates the materials and the survey used in these projects.
5.1.1 Pilot Project Procedures

Before conducting the pilot projects, it was necessary to obtain the approval of the Kent State University Institutional Review Board (KSU IRB) since this study involved human subjects. The pilot projects were conducted to ensure that the study could obtain significant data. Before the research started, permission had to be obtained from the director of LBI. After the director had approved the study to be conducted in LBI translation classes, the LBI administrative staff and the coordinator of the LBI translation classes were contacted in order to obtain their assistance in announcing the research to the respondents and in copying the flyers and the consent forms for the study. Then the first thing given to the respondents was the flyer which provides a brief explanation of this study, describes the potentially professional and intellectual benefits for the respondents, and explains the very limited risk in the study. Afterwards, they were given a consent form if they were willing to participate. The flyer and the consent form can be seen in Appendix C. After all of the consent forms were sent to me with the help of the LBI administrative staff, the respondents were sent instructions to do the tasks.

The initial pilot project was conducted in October 2013 according to the following procedures. First, the respondents were assigned a translation task (the first assignment of the pilot project – Appendix D). The participants were from two classes, an Indonesian-English translation class and an English to Indonesian translation class, and in each class they were divided into two groups in terms of different assessment models provided to them: one group was assessed using the ATA Framework and the other using the LBI Bandscale. The respondents were expected to work at their own homes individually using a computer for one hour, and they could utilize any resources to assist them, except enlisting the aid of other students and/or translators. The participants returned the completed translations directly by
email. I checked them, prepared the assessment and feedback, and returned them to the participants.

Afterwards, the respondents in both classes received their work back with the assessment and feedback, as well as the documents containing the explanation of the assessment model the participants received, two weeks after they submitted their translations. The respondents were given the final assignment instructions (Appendix D) to work on their revision, and they were allotted a week to improve their first version of the translation based on the assessment and the indirect feedback provided to them. The revision was to be sent to my email directly once they completed it. Then, they were requested to fill in a short and simple online survey (Appendix E) asking about their perspectives on the assessment they received.

I evaluated and compared the results of the respondents’ revisions from the two groups to discover whether the respondents from the first group had more or fewer errors in the revisions than the respondents from the second group (see Chapter 6). The controlling variables in the first pilot project are: (1) the time to work on the translation, and (2) the text to be translated. The independent variable was the two tools of the assessment provided to the students. The dependent variables are (1) whether the first group’s revisions have more or fewer errors than the second group’s revisions, and (2) whether the perspectives on the assessment from the respondents correlate with their translation results.

Pilot Project II was conducted in Spring 2014. As for the procedure, this second pilot project had exactly the same steps as the procedure of Pilot Project I, but the research focused on just one translation direction, from Indonesian to English. The first reason for this change is that two directions lead to more variables, and this reflects a desire to narrow the focus of the research to provide in-depth, instead of broad, results. Second, the data need to be in
English, so that the committee members will understand the results and the discussion of the study.

The translation direction into a foreign language is usually considered risky since there is a typical claim that a competent professional standard translation can only be done into the native language of the translator. However, McAlester provides good arguments that translation into a foreign language can be carried out as well as the translation into a native language depending on the condition of the country where the translators live. The following is his statement.

...in many countries where the first language of the population is not a major world language, and where there is a concomitant lack of translators who have a major world language as their language of habitual use, there exists, nevertheless, a powerful need for competent translation into such languages (McAlester, 1992), and it is desirable that such translation be done by translators who are properly trained and qualified. (McAlester, 2000:229)

From the quotation above, we can see that it is important to carry out a professional translation into a foreign language, despite concerns about absolute translation skills. Indonesian is not a major world language, and in this country the number of translation jobs into a foreign language (usually English) is as many as the number of translation jobs into the mother tongue. Many documents in business and legal situations have to be translated into English, so many administrative and clerical jobs require that a range of professionals have into-English translation skills, which is one reason people take the LBI courses.

In Pilot Project II, the respondents were divided into two groups, and they all performed the same translation task. The first group was assessed using the ATA Framework and the second group using the LBI Bandscale. The source text was the same as the one used in the first pilot project. The variables in the second pilot project are the same as the ones in the first pilot project. These pilot projects are an important and integral part of the study and the dissertation, and were designed to provide some insight on how to improve the subsequent projects.
5.1.2 Pilot Project Participants

The participants must have scored at least at the intermediate level of a language proficiency exam in order to be accepted at this translator training to ensure that their language and transfer skills are sufficient. They might vary in age, occupations, and education. In addition, they may include actual professional translators, although some might be college students or people with other professions. One common factor is that they are eager to become translators or to use translation as an ancillary skill in their professions; therefore, they register for this course, which, as previously noted, is quite expensive for people living in Jakarta, Indonesia. However, given low enrollments in each of the three terms per year, it is difficult to recruit a large enough population to produce statistically significant, scientific results. As described in Chapter 4, these results may be viewed as speculative or indicative, but they are expected to still be useful for translation assessment development (see Muñoz and Conde, 2007).

In the pilot project in the Fall 2013, the participants were two (2) students out of five (5) registered in the English to Indonesian translation class, and all the 13 people registered in the Indonesian to English translation class. One participant from the Indonesian to English translation class withdrew from the project by not submitting the revision or the final assignment. All respondents in the Indonesian-English translation class were very enthusiastic to be participating in the research. They were grateful and happy to have their translations assessed and given feedback. As a result, almost all of them returned their first and final assignments and filled in the survey, and there was only one respondent who did not return the revision. Actually, the students of LBI translation classes rarely or almost never receive feedback for their work in writing. All of the teachers just give exercises in class, and
the results are always discussed in class too. When the students are given homework to do a translation, most of them will never do it, and the reason is because they are busy with their jobs.

In the second pilot project in Spring 2014, the focus of the research was narrowed to only the participants from the Indonesian to English translation class. In Pilot Project II, there were only five participants who were willing to take part in this project out of seven students. Unfortunately, out of five, there were only two who submitted the final assignments or the revisions and filled in the survey. With very few respondents in this term, the results cannot be compared and analyzed, as the respondents returning the final assignment were those whose translations were assessed using the ATA Framework, and there is no data for the translations assessed using the LBI Bandscale in this term.

5.1.3 Pilot Project Materials and Survey

As described in the procedure, before the respondents were provided with task instructions, they were given a flyer and a consent form (see Appendix C), so these two documents are also important materials for the study. The instructions for the first assignment include the source text. The English source text given to the respondents in the first pilot project is a non-technical text consisting of 170 words. The title of the text is “Eurozone Unemployment Reaches New Record High in April.” It is about the high unemployment rate in Europe, and it was taken from a BBC news website (http://www.bbc.co.uk/news/business-22727373). The Indonesian source text in this project is also a non-technical text which consists of 182 words. The title is “Tembus Rekor Lagi, Penggangguran di 17 Negara Pengguna Euro Capai 19,38 Juta” (Breaking the Record Again, the Unemployment Rate in 17 Euro Countries Reaches 19.38 Millions). It was taken from an Indonesian news website
and it is also about the increase in the unemployment rate in Euro countries. Both English and Indonesian source texts have quite similar content with some differences. The respondents were asked to translate only the highlighted paragraphs in both source texts. In the second pilot project in Spring 2014, as the study focused on the Indonesian-English translation class respondents, the source text used was only the Indonesian source text, which is exactly the same as the one used in the first pilot project. The two source texts can be seen in Appendix D.

The survey was conducted using the Survey Monkey website. The title of the survey used for the pilot projects is “Perspectives on Translation Assessment.” It can be found at http://www.surveymonkey.com/r/GRGJZG8. There are six (6) questions. Four questions are closed-ended questions asking respondents to choose one suitable answer from the answers provided, and two questions are open-ended. According to Peterson, in open-ended questions the participants are free to provide any answer in their own opinions, while in closed-ended questions answers are provided for the participants to choose (2000:29). The details of the survey can be seen in Appendix E.

5.2 The Formal or Second Projects

These second projects are the formal projects of the study. There are several changes in these projects based on the results of the pilot projects. This subchapter is divided into three parts. The first part discusses the procedures that are different from the procedures of the pilot projects, and there is an explanation of using the independent judge as Rater 2 in these projects. The second part describes the participants in the second projects, which were
carried out in the Summer 2014, the Fall 2014, and the Spring 2015. The last part explains the materials and the survey used in the second projects.

5.2.1 Second Project Procedures

The second round of projects was conducted in the Summer and Fall 2014, and continued in Spring 2015 in order to gather additional data. In these two projects, the translation direction is also only from Indonesian to English. Furthermore, there were several changes in the procedure, in the text to be translated, and in the survey. An independent judge was involved in order to avoid biased results in the final assessments of the respondents’ corrected translations.

First, after having been given the flyer and the consent form, the participants were asked to translate two paragraphs (the first assignment – Appendix D), and one paragraph was assessed using the ATA Framework and the other using the LBI Bandscale. This change in the project design enabled the participants to see and compare directly the two assessment models. The participants in the pilot projects could only see one assessment model, and they did not know about the other one. Given the fact that they are unfamiliar with formative assessment in any form, they have no standard of comparison between the two forms. The first group in particular was so pleased to have feedback of any kind that they were undistinguished in their enthusiasm for the project.

The participants in these projects were given three (3) hours to do the translation instead of only one hour because the text is a little bit longer than the one used in the pilot projects. It should be noted that the time frames allowed are generous and that the students should be able to complete the tasks comfortably without feeling rushed. Nevertheless, time should not be a determining variable in this study; however, the duration for doing the task
must be mentioned as part of IRB requirements. Similar to the previous projects, the participants received the assessment and feedback for their translations, as well as all the documents explaining the ATA Framework and the LBI Bandscale, within two weeks, and they were asked to do the revision within one week. Afterwards, they were asked to fill in a short and simple survey as well, but the survey is different from the one in the previous projects, although the theme of the survey remains the same, which is to reveal their perspectives on the assessment models provided for them.

Another change in the research design is the involvement of an independent judge to also evaluate the respondents’ revisions in order to avoid biases and to produce more “objective” results. The independent judge is a professiona l translator and a translation academician from Indonesia, demonstrates a recognized mastery of both English and Indonesian, and is an experienced evaluator of student translations. The independent judge was asked to assess the final translations of the respondents by marking errors and giving a grade (A, B, C, or D). Because he has been the LBI translation class coordinator since Spring 2014 and he is not familiar with the ATA Framework, he applied the LBI Bandscale to assess the revisions. All the revisions sent to him were made anonymous. The results from the independent judge (Rater 2) will be correlated with the results from Rater 1 (me) to reveal whether there is good inter-rater reliability between the two raters (see Chapter 6).

5.2.2 Second Project Participants

For the second project in the Summer 2014, there were seven students enrolled, but there were only five participants who were willing to take part. Four out of five submitted all assignments, and those four respondents filled in the survey as well. As for the project in the Fall 2014, there were ten (10) participants enrolled to the course, and there were eight (8)
who were willing to take part in this research. Out of the eight respondents, there were six (6) who returned the final assignments, but there were only five (5) who filled in the survey. In the Spring 2015, there were six participants enrolled to the course, but there were only three (3) who took part in the study. They all completed the first and the final assignments, but only two filled in the survey. It is possible that several respondents were not aware of the importance of completing all three tasks in this study. Future studies of this nature could perhaps include a rationale for filling in the survey or perhaps some incentive to encourage students to fill in the survey.

5.2.3 Second Project Materials and Survey

The flyer and the consent form for these second projects are not much different from the ones in the pilot projects. Only the title was changed, but the topic remains basically the same. The source text used in the second projects is also a non-technical text, but it consists of two paragraphs with the total of 314 words. This first paragraph is about the high tuition fee for a special international school in Indonesia named *Rintisan Sekolah Bertaraf Internasional* (Pioneered International Standard School) or RSBI for short, and it consists of 183 words. The second paragraph is about the use of English as the language of instructions in RSBI in Indonesia, and it consists of 131 words. The second project materials can be seen in Appendices C (the flyer and the consent form) and D (the first and final assignments).

The survey for the second projects was also conducted using the Survey Monkey website. It consists of seven simple questions. The title of the survey is “Comparing Two Translation Assessment Models.” The first two questions are closed-ended questions with two answers provided. Four questions are also closed-ended questions with a Likert survey format with five answers provided in each question, and only one answer is required. A
The Likert scale is a psychometric scale and the most widely used approach to scaling responses in survey research, and this scale is named after its inventor, Rensis Likert, a psychologist (http://poincare.matf.bg.ac.rs/~kristina/topic-dane-likert.pdf). The last question is an open-ended question. The survey can be found at https://www.surveymonkey.com/r/J6LHRTL, and it can be seen in Appendix E.

This chapter has described in detail the methodology applied for this study for both the pilot projects and the second projects. From the explanation of the methodology, it can be observed that the projects involved several tasks for the respondents: (1) the first assignment, which is creating a draft translation, (2) the final assignment, which is revising the first assignment based on the assessment and feedback given, and (3) filling in an online survey. The procedure asks the respondents to do quite a lot, so it is not surprising that many students of the LBI translation classes were not willing to participate in the research, and several of them who were willing did not complete all assignments. Those who completed all tasks were eager learners of translation, and they were happy to receive feedback on their translations. Some of them implicitly stated in their email that they hardly ever had their translations assessed thoroughly with established translation assessment models like the LBI Bandscale and the ATA Framework.

It can also be seen that the number of students enrolled to the LBI translation classes is relatively small in every term. The highest number is never more than 15 students, and the range is usually between 5 to 10 students. This contributes to the small number of the respondents in this study, although the data collection was carried out for five terms or almost two years. To have a bigger population of the respondents from the LBI translation classes to generate significant and scientific results, the data should be collected over more than five years, which in this case is highly impossible due to the time constraints imposed on writing
This dissertation. It will be more feasible to conduct this kind of longitudinal study when I return to my country to teach.

Furthermore, it can be seen that I, the researcher, was not the teacher of the class, and communication was conducted via email. Thus, there was no two-way discussion between me and the respondents. In an ideal feedback situation, students should be trained to be familiar with the assessment model(s) provided for them, and the class should also be conducted in a student-centered approach where the students can critique each others’ translations so that they have a rich and varied amount of feedback, besides having seen several options for successful or functional translation. Although the respondents were welcomed to ask any questions in relation to the tasks or the research, none of them posed any. They might have felt reluctant, as I was far away from them and they have never met me in person. As a result, even though this study ideally supports the student-centered approach and the results from the study should be implemented in a learner-oriented classroom, the condition of the research cannot implement either classroom approach, unfortunately. In the next chapter, Chapter 6, the definitions of translation errors and the explanation of error typology will be explained. Then, the results of the three tasks from the pilot projects and the second projects will be analyzed and discussed.
CHAPTER 6

RESULT ANALYSIS AND DISCUSSION

This chapter analyzes and discusses the results from all the projects conducted. Since most of the data from this study refer to translation errors, their definitions and typology must be explained first before continuing to the discussion of the project results. Besides translation errors, there are also data from survey results which reflect the respondents’ perspectives. The inter-rater reliability between Rater 1 and Rater 2 in assessing the revisions of the respondents will be described as well to reveal the validity and reliability of the raters. The analyses of the pilot projects and the formal projects will be treated separately as they have different designs. The weaknesses of the pilot projects will be explained because these provide the reasons for the changes made in the formal or second projects.

6.1 Translation Errors

This subchapter defines translation errors based on trends in Translation Studies. In regards to all the definitions proposed by several TS scholars, a definition of translation errors based on the projects will be proposed. Then the types of translation errors from several translation assessment models and TS scholars will be elaborated with examples. The translation error typology based on these projects will be generated as well following the discussion of error types from many TS scholars.
6.1.1 Definitions of Translation Errors

Translators, both novices and professionals, make errors (Séguinot, 1990:68) because humans have limited cognitive processing capacity or limitations on short-term memory (1989:75), and because translators sometimes have vocabulary and knowledge gaps that are not always filled in time (1990:68). Therefore, translation errors can be found almost in any translation, especially in the first draft. Nevertheless, what are translation errors? Are they different from language errors? Vinay and Dalbernet are among the earliest TS scholars who discuss translation errors, although they do not explicitly provide a definition of translation errors per se. They state that translation errors occur when translators do not carefully pay attention to the subtle differences of the meanings of words that on the surface appear to be interchangeable (1958/1995:58).

Viewed from the standpoint of equivalence between a source text (ST) and a target text (TT), a translation error, according to Koller, is considered as non-equivalence between ST and TT or non-adequacy of the TT (1979:216). Viewed from a functionalist approach, Sigrid Kupsch-Losereit (1985) was the first to introduce the notion of translation errors. Her definition of a translation error is as “an offence against: (1) the function of the translation, (2) the coherence of the text, (3) the text type or text form, (4) linguistic conventions, (5) culture- and situation-specific conventions and conditions, (6) the language system” (1985:172). Williams provides a definition for a major translation error but not for a minor translation error. He explains that a major translation error is “the complete failure to render the meaning of a word or group of words conveying an essential part of the message of the document” (1989:24). Nord states that translation errors happen due to structural distinctions in the syntax and suprasegmental features of the two languages (1997:66). Based on the
skopos theory or functionalism, the definition of a translation error is stated “as a failure to carry out the instructions implied in the translation brief and as an inadequate solution to a translation problem” (75) and is the relative fulfillment of the TT-function and the receiver’s expectations (Schmitt 1998:394; Nord 2009:190).

Thus, based on these definitions, what a translation error is has a variety of meanings depending on translation theories and norms (Hansen, 2010:385). Hansen also states that translation errors appear since something “goes wrong” during the transfer and movement from the ST to the TT (Ibid.). Then, Conde discusses the differences between language errors and translation errors. He maintains that language errors are found in the TT and are often similar to the errors on target language expression consisting of mistakes in vocabulary, syntax, grammar, punctuation, coherence, style, etc., whereas translation errors are “explained by the existence of a previous text: the source text upon which the target text depends” (2013:98). Both Williams (1989) and Conde (2013) classify translation errors as errors of meaning. Lee and Ronowick also explain the differences between translation errors and language errors, which are similarly described by Conde (2013). They state that “translation errors, regardless of whether they are major or minor, are associated more with the source text, while language errors are associated more with the target text” (Lee & Ronowick, 2014:42).

One of the assessment models in these projects, the ATA Framework, provides a simple definition of a translation error. It is written in the form with the title ATA Framework for Error Marking (see Appendix B) that translation errors are negative impact(s) on the understanding or use of a target text. Translation errors in this model are also called strategic or transfer errors. The other model used in this study, the LBI Bandscale, however, does not provide any definition of translation errors. Based on the projects of this study, translation
errors can be classified as mistranslation, misunderstanding, and unclear, awkward, and ambiguous meanings and/or expressions.

6.1.2 Translation Error Typology

Several Translation Studies scholars and translation assessment models have proposed their own error typologies. Chien states that up to now there have not been commonly agreed or confirmed categories of translation errors (2015:91). In other words, there is still no universal translation error typology. This might be the case because, first, the definitions of translation errors vary according to different translation theories, which leads to different categorizations of errors as well. Second, different translation language pairs could result in different types of errors. In the following we can see the different error typology proposed, including the error typology from these projects.

Williams divides errors into translation/transfer errors and language errors, and each consists of major and minor errors (1989, 2001). This error typology has been applied by SICAL (the Canadian Language Quality Measurement System) developed by Canadian government’s Translation Bureau where Williams was the head of the committee that designed the assessment model and that decided the categories of errors (Ibid.). Williams explains that there are three types of errors or defects in industrial quality control theory: critical, major, and minor (1989:23). A critical error is defined as an error that can lead to dangerous or hazardous conditions for anyone applying the product with such an error. A major error is defined as an error resulting in failure or reducing the usability of the product with such an error. A minor error is defined as an error that will not reduce the usability of the product with such an error; in other words, this error will have only little impact on the effectiveness of a tool or a product (Ibid.).
Major translation errors, according to Williams, are the combination of critical and major errors as described for industry (24) and involve macro-level misinterpretations (2001:331). The examples of major translation errors of SICAL described by Lee & Ronowick are serious mistranslation, significant omissions, and nonsense (2014:43). In contrast, minor translation errors are micro-level misinterpretations (Williams, 2001:331), which include [trivial] mistranslations, shift in meaning, ambiguity, addition, and omission (2001:328-333; Lee & Ronowick, 2014:43). Furthermore, a major language error refers to a serious language error, such as an unwarranted neologism (Williams, 1989:24-25), unintelligible language, or grossly incorrect language (Lee & Ronowick, 2014:43). Minor language errors are small errors, in word-level structures, which do not disrupt the meaning in the target text (Williams, 2001:331). His examples of minor language errors involve diction, punctuation, syntax, style, morphology, cohesion devices, spelling, and others (*Ibid.*; Lee & Ronowick, 2014:43). Unfortunately, Williams and Lee & Ronowich do not provide definitions and examples of those types of errors. The table below shows the SICAL error typology.

Table 6.1
The Canadian Language Quality Measurement System (SICAL)

<table>
<thead>
<tr>
<th>Major error</th>
<th>Minor error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Translation error</td>
<td>Serious mistranslation</td>
</tr>
<tr>
<td></td>
<td>Significant omission</td>
</tr>
<tr>
<td></td>
<td>Nonsense</td>
</tr>
<tr>
<td></td>
<td>[Trivial] Mistranslation, Shift in Meaning</td>
</tr>
<tr>
<td></td>
<td>Ambiguity, Addition, Omission</td>
</tr>
<tr>
<td>Language error</td>
<td>Unintelligible language</td>
</tr>
<tr>
<td></td>
<td>Grossly incorrect language</td>
</tr>
<tr>
<td></td>
<td>Unacceptable neologism</td>
</tr>
<tr>
<td></td>
<td>Diction</td>
</tr>
<tr>
<td></td>
<td>Punctuation</td>
</tr>
<tr>
<td></td>
<td>Syntax, Style, Morphology</td>
</tr>
<tr>
<td></td>
<td>Cohesion Devices</td>
</tr>
<tr>
<td></td>
<td>Spelling</td>
</tr>
<tr>
<td></td>
<td>Others</td>
</tr>
</tbody>
</table>

Pym has proposed two types of translation errors: binary errors and non-binary errors. He explains that “a binary error opposes a wrong answer to the right answer,” so for this type of errors it is about “right” and “wrong” (1992:282). Binary errors are typically very rapidly and punctually corrected (285). On the contrary, non-binary errors or non-binarism “requires that the target text actually selected be opposed to at least one further target text, which could also have been selected, and then to possible wrong answers” (282). There is a minimum of two right answers and two wrong answers for non-binary errors (Ibid.). The time needed for correcting non-binary errors can take long until there are no more significant differences (285). Pym maintains that all translational errors (as opposed to language errors) are non-binary by definition (Pym’s definition), but non-binary errors are not necessarily all translational (283).

For pedagogical purposes, Nord categorizes translation errors into four types: pragmatic, cultural, linguistic, and text-specific (1997:64). Pragmatic translation errors occur because of insufficient solutions to translation-oriented understanding and resolution of pragmatic ambiguities posed by the source text (75). An example of a pragmatic translation problem is a lack of receiver orientation. According to Nord, this problem appears because there are differences between the source text and target text situations, and it can be identified by observing extratextual factors, such as sender, receiver, medium, time, place, motive, text function (65). For instance, when we translate a legal term from Indonesian to English, such as the expression *putusan sela*, the translation tends to be ‘(court) decision’ as it is the literal translation of the term; however, it should be translated as ‘order’ because functionally it is the meaning. Thus, to avoid this type of pragmatic translation errors, we must pay attention to the function of a term or an expression in the target text. The second type of translation errors is cultural translation errors. These errors occur “due to an inadequate decision with regard to reproduction or adaptation of cultural-specific conventions” (75). Because conventions from
one culture to another are not the same, cultural translation errors will be different for
different language pairs (66).

The third type of translation errors is linguistic translation errors. These errors are
causd by inadequate translation when the focus is on language structures, such as in foreign
language classes (75). Linguistic translation errors are restricted to language pairs since
structural differences in the syntax and suprasegmental features of one language pair might
be different from those of another language pair (66). The fourth type of translation errors is
text-specific translation errors. These errors, of course, depend on the text translated and are
typically evaluated from a functional or pragmatic point of view (76). Solutions to these text-
specific errors can never be generalized and cannot even be applied to similar cases. The
eamples of these errors are incorrect figures of speech, improperly formed neologisms or
puns, etc. (67).

Hansen categorizes translation errors into pragmatic, text linguistic, semantic,
iomatic, stylistic, morphological, and syntactical errors (2009, 2010). Pragmatic errors refer
to misinterpretation of the translation brief and/or the communication situation (2009:320).
The examples of these errors are an incorrect translation type, lack of important information,
warranted omission of ST units, too much information related to the ST and/or the TT
receiver’s needs in the situation, disregarding norms and conventions as to genre, style,
register, abbreviations, etc. (Ibid.). Text-linguistic errors constitute a violation of the
semantic, logical, or stylistic coherence, such as incoherent text caused by incorrect
connectors or particles, incorrect or vague reference to phenomena, unclear temporal
cohesion, incorrect category (using active voice instead of passive voice, or the other way
around), incorrect modality, incorrect information structure caused by word order problems,
and unmotivated change of style (320-321).
Semantic (lexical) errors are incorrect choices of words or phrases (321). Idiomatic errors are semantically correct words and phrases, but these errors will not be used in an analogous context in the TL. Stylistic errors refer to incorrect choices of stylistic level, stylistic elements, and stylistic devices. Morphological errors, also called “morpho-syntactical errors,” include incorrect word structure or mistakes in number, gender, or case. Syntactical errors include incorrect sentence structure, for example (*Ibid.*).

Angelone describes several translation errors which are quite similar to Hansen’s list, but there are a few which are different. His study focuses on lexical errors, syntactic errors, stylistic errors, and mistranslation errors (2013:263). Lexical errors refer to using incorrect terminology, false cognates, and incorrect or weak collocations. Syntactic errors include errors in word order, verb tense, incomplete or run-on sentences, and subject-verb agreement. Stylistic errors are caused by inappropriate register, lexical or grammatical inconsistency, or lexical and grammatical redundancy. Mistranslation errors happen due to inappropriate additions or omissions and errors causing incorrect meaning transfer to appear (*Ibid.*).

Lee & Ronowick have their own categories of translation errors based on the Korean-English translation. Their framework of error typology includes the causes of errors, the types of errors, the results of errors, and the significance of errors (2014:47). Below is the table containing the framework or error typology from Lee & Ronowick.

### Table 6.2
Translation Error Typology of Lee & Ronowick (Korean to English)

<table>
<thead>
<tr>
<th>Causes of errors</th>
<th>Miscomprehension of source text</th>
</tr>
</thead>
<tbody>
<tr>
<td>Types of errors</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Misuse of Korean</td>
</tr>
<tr>
<td>Lexical errors</td>
<td>Incorrect word</td>
</tr>
<tr>
<td></td>
<td>Loan word</td>
</tr>
<tr>
<td></td>
<td>Word to be refined</td>
</tr>
<tr>
<td></td>
<td>Redundant word</td>
</tr>
<tr>
<td></td>
<td>Incorrect terminology</td>
</tr>
<tr>
<td></td>
<td>Collocation</td>
</tr>
<tr>
<td></td>
<td>Parts of speech</td>
</tr>
<tr>
<td></td>
<td>Ending</td>
</tr>
</tbody>
</table>

89
Lee & Ronowick explain that the ‘causes of errors’ category aims to identify which of the two languages is the source of errors in the translations. They describe in their framework that the causes of errors derive from errors in the comprehension and reformulation phases (Ibid.). More specifically, in the Korean-English translation, the causes of errors are from the misinterpretation of the source text, such as misunderstanding of words, phrases, and clauses of the source text (48-49). ‘Types of errors’ refer to the skills and knowledge of students that must be improved, and there are three types of errors in this study: lexical (incorrect word, loan word, word to be refined, redundant word, incorrect terminology, collocation), syntactical (parts of speech, ending, voice, word order, agreement, incomplete sentence, tautology, omission), and ‘hygiene’ (spacing, punctuation) errors (Ibid.).

Lee & Ronowick explain that ‘results of errors’ involve the impacts resulting from the different types of errors (48). These results of errors have three aspects: (1) fidelity in meaning transfer from the source text to the target text, such as distortion and ambiguity, (2) grammar in the target text, such as TT unacceptability, (3) the coverage of content, such as information loss. Then, ‘significance of errors’ in their study is adopted from the NAATI and SICAL methods of assessment, and this aims to decide if translations are good or bad. This significance of errors consists of major and minor errors, and the concept of these major and

---

**Table: Types and Results of Errors**

<table>
<thead>
<tr>
<th>Source: Lee &amp; Ronowick, 2014:49</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Results of errors</strong></td>
</tr>
<tr>
<td><strong>Significance of errors</strong></td>
</tr>
</tbody>
</table>
minor errors follows the concept discussed by SICAL and NAATI. They further explain that minor errors include easily corrected errors in the target text without analysing the source text and simple language errors. Major errors include errors that require analysis of the source text and errors that cause the target text to be incomprehensible (Ibid.).

In the ATA Framework (see Appendix B), the translation errors are not divided into major and minor. Instead, the errors are categorized into translation/strategic/transfer errors and mechanical errors and weighted according to an exponential scale of 1/2/4/8/16, which takes the place of the notion of minor and major. When an error type is given the value 1/2/4, it means the error is considered minor. When it receives a value 8 or 16, it means the error is major. As described above, translation errors in this model refer to negative impact(s) on understanding or use of a target text. These errors include mistranslation (MT), misunderstanding (MU), addition (A), omission (O), terminology or word choice (T), register (R), faithfulness (F), literalness (L), faux ami (FA – false friend), cohesion (COH), ambiguity (AMB), style (ST – inappropriate for specified type of text), and other (OTH – if there are other types of errors that have not been covered in these translation errors). Mechanical errors in this model refer to negative impact(s) on overall quality of a target text. These errors include grammar (G), syntax (SYN – phrase/clause/sentence structure), punctuation (P), spelling or character (SP/CH), diacritical marks or accents (D), capitalization (C), word form or part of speech (WF/PS), usage (U), and others (OTH – if there are other types of mechanical errors that have not been covered). The points for G, SYN, P, WF/PS, U, and OTH can be 1, 2, or 4, and the points for SP/CH, D, and C can be 1 or 2. The rational for the scale is that the reader of the target text will likely notice the error, perhaps be irritated by it, but will not misread the message of the test. The explanation of each error can be seen in Appendix B.
In contrast, the LBI Bandscale (see Appendix A) has no list of errors. It presents a set of model scales, each of which focuses on the translator’s understanding or comprehension of the source text (ST) and on the translator’s rendition of the target text (TT). There are several types or error mentioned, but not described or explained in this model, such as spelling errors, punctuation errors, incorrect terminology or word choices, collocation errors, and misunderstandings. Nor does this model categorize errors into major and minor, and there are no points to scale the errors mentioned. Each grade provided (A, B, C, and D) refers to overall performance, the strengths and the weaknesses of a translation result.

In this study, even though errors are not equal, each error is scaled or counted as one error since this observational study does not focus on types of errors; instead, it emphasizes the decrease in the number of translation errors in the revisions and the effectiveness of the two assessment models as formative assessment in improving student performance. In these projects where the respondents were asked to translate from Indonesian to English, a number of error types were discovered. The errors include incorrect terminology, incorrect usage, incorrect syntax, mistranslation or misunderstanding, literalness/faithfulness, ambiguity, omission, addition, incorrect capitalization, incorrect punctuation, incorrect spelling, grammar errors, incorrect word form, and incorrect word order. The following table shows the errors found in the translations and revisions produced by the respondents in these projects plus the frequency of the error types.

<table>
<thead>
<tr>
<th>Types of Errors</th>
<th>Frequency of Errors</th>
<th>Decrease of Errors</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Translation Errors (TE)</td>
<td>Revision Errors (RE)</td>
</tr>
<tr>
<td>Incorrect terminology</td>
<td>39</td>
<td>11</td>
</tr>
<tr>
<td>Incorrect usage</td>
<td>280</td>
<td>130</td>
</tr>
<tr>
<td>Incorrect syntax</td>
<td>60</td>
<td>16</td>
</tr>
<tr>
<td>Mistranslation/</td>
<td>23</td>
<td>9</td>
</tr>
</tbody>
</table>

Table 6.3
Translation Errors based on This Study (Indonesian to English)
Based on the table and the figure above, the most common errors found in the projects are grammatical errors. These errors include singular-plural and countable-uncountable errors (e.g., a lot of book), incorrect tenses, incorrect articles, incorrect prepositions, incorrect
connectors, problems with subject-verb agreement (e.g., ‘he study’), and problems with active-passive voice. The second most common errors are usage errors, including awkward expressions, incorrect words or expressions, unclear expressions, and inappropriate expressions. An example of this incorrect usage is the expression ‘the number of unemployment,’ which should be ‘the number of unemployed’ or ‘unemployment rate.’ The third most common errors are omission errors which occur when there are expressions, sentences, words, articles, and prepositions missing or not translated into the target text. For instance, in a dependent clause ‘….of the area in which RSBI located,’ there is a verb (‘is’) missing, so it should be ‘in which RSBI is located.’ These kinds of errors reflect the fact that the students are translating into their Language B. One might anticipate in a future study involving translation into Indonesian from English that there might be fewer target language grammatical and stylistic errors.

The fourth most common errors found are incorrect punctuations. For example, in English we must use a comma for the number above a thousand, such as 1,200, and not a period (e.g., 1.200) as in Indonesian. The fifth most common errors are word form errors that involve incorrect forms of a word. For example, a word should be a verb, but it is written in a noun form (e.g., ‘to strength’ instead of ‘to strengthen’). The sixth most common errors are syntax errors which include run-on-sentences and incorrect sentence structure. For example, in a dependent clause ‘….if Indonesian values to be instilled,’ this clause has no subject and no verb, and it is more like a phrase than a clause. The seventh most common errors are addition errors that happen when there are unnecessary expressions, sentences, words, articles, prepositions added to the translation, and when there are redundants. For instance, in the title ‘News about the high cost of school’s tuition,’ the expression ‘cost of’ is redundant based on the source text because there is ‘school’s tuition,’ so it should be ‘News about the high school’s tuition.’
The eight most common errors are terminology errors. For example is the incorrect full form for the abbreviation EUROSTAT. The correct expansion for EUROSTAT is Statistical Office of European Communities; however, several respondents wrote ‘European Union Statistic’s Agency’ or ‘Europe Statistic Commission Office.’ The ninth most common errors are spelling errors involving incorrect spelling. For instance, the word ‘language’ was written ‘languange’ by one of the respondents, the word ‘quite’ was written ‘quiet.’ or the word ‘economist’ was written ‘ekonomist.’ The tenth most common errors are mistranslation or misunderstanding errors which reflect the change of the meaning or the introduction of different meanings in the target text from the source text. For example, in the source text of the pilot projects, it is mentioned that ‘for this month the unemployment rate reached 12.2%, while in the previous month it was 10%, so the increase is 2.2%.’ However, several respondents misunderstood the text, and they translated that ‘the increase was 10% from the previous month.’

The eleventh most common errors are capitalization errors. These errors mean that a letter (usually the first letter) in a word should be in a capital letter, but it is not, or the other way around. For instance, one of the respondents wrote: ‘however, it was….’ The word ‘however’ is located at the beginning of a sentence, so it should be ‘However, it was….’ The next error types are the ones which do not frequently occur or only appear few times. The twelfth most common errors or the third least common errors are faithfulness or literalness errors that refer to the expressions in the target text being translated very close to the meaning and to the structure of the expressions in the source text, so the translation becomes awkward or the meaning is not transferred correctly. For example, the expression ‘17 Euro users’ is a literal translation from the Indonesian source text, while the appropriate translation should be ‘17 Euro country users.’ The second least common errors are ambiguity errors that include words and expressions having ambiguous or unclear meanings. For instance, the sentence ‘the
The lowest number of unemployment is Austria with 4.9%’ is confusing, as it should be ‘Austria has the lowest number of unemployed with 4.9%.’ The least common errors are word order errors that involve putting a word or an expression in an uncommon or inappropriate location in a sentence, although it might not interrupt the meaning of the whole sentence or the entire target text. For instance, a time phrase ‘in April 2013’ was put in the middle of a sentence, while it should be at the end of the sentence.

6.2 Pilot Project Result Analysis and Discussion

In this subchapter, when the data from the pilot projects are analyzed manually (using Microsoft Excel), the results reveal the mean or the average of the error difference between assessment using the ATA Framework and assessment using the LBI Bandscale. The statistical calculations with t-test using SPSS (Statistical Package for the Social Sciences) will show whether there is a significant difference or not between the improved translation results assessed using the ATA Framework and the ones assessed using the LBI Bandscale. Then, several weaknesses from the pilot projects and what changes that had to be made for the second projects will be elaborated.

6.2.1 Quantitative Results and Interpretation of the Pilot Projects

As described in Chapter 5, the pilot projects were conducted in Fall 2013 and Spring 2014. However, since the results from the Spring 2014 are not sufficient due to only two respondents returning their final assignments, just the results from the Fall 2013 term can be analyzed and discussed. The following table contains the data from the pilot projects in the Fall 2013 and Spring 2014.
## Table 6.4
Pilot Project Data

### Pilot Project Results in Fall 2013 with the ATA Framework

<table>
<thead>
<tr>
<th>Respondents</th>
<th>Translation Errors (TE)</th>
<th>Revision Errors (RE)</th>
<th>Error Difference (ED)</th>
<th>Percentage of Error Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fall2013ATA</td>
<td>20</td>
<td>3</td>
<td>17</td>
<td>85.00</td>
</tr>
<tr>
<td>Fall22013ATA</td>
<td>51</td>
<td>7</td>
<td>44</td>
<td>86.27</td>
</tr>
<tr>
<td>Fall32013ATA</td>
<td>15</td>
<td>2</td>
<td>13</td>
<td>86.67</td>
</tr>
<tr>
<td>Fall42013ATA</td>
<td>35</td>
<td>20</td>
<td>15</td>
<td>42.86</td>
</tr>
<tr>
<td>Fall52013ATA</td>
<td>33</td>
<td>18</td>
<td>15</td>
<td>45.45</td>
</tr>
<tr>
<td>Fall62013ATA</td>
<td>26</td>
<td>17</td>
<td>9</td>
<td>34.62</td>
</tr>
<tr>
<td>Fall72013ATA</td>
<td>39</td>
<td>21</td>
<td>18</td>
<td>46.15</td>
</tr>
<tr>
<td><strong>Average/Mean</strong></td>
<td><strong>31.29</strong></td>
<td><strong>12.57</strong></td>
<td><strong>18.71</strong></td>
<td><strong>61.00</strong></td>
</tr>
</tbody>
</table>

### Pilot Project Results in Fall 2013 with the LBI Bandscale

<table>
<thead>
<tr>
<th>Respondents</th>
<th>Translation Errors (TE)</th>
<th>Revision Errors (RE)</th>
<th>Error Difference (ED)</th>
<th>Percentage of Error Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fall2013LBI</td>
<td>36</td>
<td>25</td>
<td>11</td>
<td>30.56</td>
</tr>
<tr>
<td>Fall22013LBI</td>
<td>28</td>
<td></td>
<td></td>
<td>0.00</td>
</tr>
<tr>
<td>Fall32013LBI</td>
<td>27</td>
<td>13</td>
<td>14</td>
<td>51.85</td>
</tr>
<tr>
<td>Fall42013LBI</td>
<td>31</td>
<td>17</td>
<td>14</td>
<td>45.16</td>
</tr>
<tr>
<td>Fall52013LBI</td>
<td>43</td>
<td>16</td>
<td>27</td>
<td>62.79</td>
</tr>
<tr>
<td>Fall62013LBI</td>
<td>54</td>
<td>29</td>
<td>25</td>
<td>46.30</td>
</tr>
<tr>
<td><strong>Average/Mean</strong></td>
<td><strong>36.5</strong></td>
<td><strong>20</strong></td>
<td><strong>18.2</strong></td>
<td><strong>47.33</strong></td>
</tr>
</tbody>
</table>

### Pilot Project Results in Spring 2014 with the ATA Framework

<table>
<thead>
<tr>
<th>Respondents</th>
<th>Translation Errors (TE)</th>
<th>Revision Errors (RE)</th>
<th>Error Difference (ED)</th>
<th>Percentage of Error Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spr2014ATA</td>
<td>48</td>
<td>5</td>
<td>43</td>
<td>89.58</td>
</tr>
<tr>
<td>Spr22014ATA</td>
<td>22</td>
<td>3</td>
<td>19</td>
<td>86.36</td>
</tr>
<tr>
<td>Spr32014ATA</td>
<td>45</td>
<td></td>
<td></td>
<td>0</td>
</tr>
<tr>
<td><strong>Average/Mean</strong></td>
<td><strong>38.33</strong></td>
<td><strong>4.00</strong></td>
<td><strong>31.00</strong></td>
<td><strong>87.97</strong></td>
</tr>
</tbody>
</table>
Based on the data from Fall 2013, ideally there should have been seven respondents assessed using the ATA Framework and six respondents assessed using the LBI Bandscale. However, one respondent assessed using the LBI Bandscale never returned the final assignment, so there are just five results assessed using the LBI Bandscale. To determine whether one of the two translation assessment models is more effective in reducing errors in the revision, the number of translation errors (TE) is substracted from the number of revision errors (RE), and this generates the error difference (ED).

\[
\text{TE} - \text{RE} = \text{ED}
\]

The first statistical calculation (Table 6.4) is designed to yield the mean of ED. The mean, according to Heiman, is “the score located at the exact mathematical center of a distribution” (2001:170), and most people also call this as the average. The average/mean of ED for the ATA Framework in this project is 18.71 or 61%, whereas the mean of ED for the LBI Bandscale is 18.2 or 47.3%. According to this calculation, it seems that there is a difference in the effectiveness of these two assessment models, and that the ATA Framework is more effective than the LBI Bandscale.

However, in the second statistical calculation conducted with SPSS (Statistical Package for the Social Sciences), there is a different result from the one in the above
paragraph. The second calculation involves the $t$-test. The following is the definition of the $t$-test.

The $t$-test is for testing a single-sample mean when (a) there is one random sample of interval or ratio data, (b) the raw score population is normally distributed, and (c) the standard deviation of the raw score population is estimated by computing $s_x$ [standard deviation] from the sample data. (Heiman, 2001:393)

The symbol for the answer obtained from the $t$-test is $t_{obt}$, and the critical (significant) value of $t$ is symbolized by $t_{crit}$ (369), which shows the boundary of the significant value of a sample mean. If $t_{obt}$ is beyond $t_{crit}$, it means that the sample mean is in the region of rejection, and the null hypothesis can be rejected (370). If $t_{obt}$ is not beyond $t_{crit}$, it means that the sample mean does not lie in the region of rejection, and the null hypothesis cannot be rejected. The region of rejection is “the part of a sampling distribution containing values that are so unlikely to occur that we ‘reject’ that they represent the underlying raw score population” (325).

To make this concept clearer, the following two graphs contain models for sampling distribution that show the position if $t_{obt}$ is beyond $t_{crit}$ (Figure 6.2) and if $t_{obt}$ is not beyond $t_{crit}$ (Figure 6.3).
In Figure 6.2 above, the $t_{obt}$ lies in the region of rejection or is beyond the $t_{crit}$, so it means the sample mean is significant, and the null hypothesis can be rejected.

In Figure 6.3 above, the $t_{obt}$ does not lie in the region of rejection or is not beyond the $t_{crit}$, so it means the sample mean is not significant, and the null hypothesis cannot be rejected.

The design for these pilot projects is a *between-subjects design*; in other words, it is a *design having independent samples*. A between-subjects design is a design where
“participants are selected for each condition without regard for who is selected in another condition, and each participant serves in only one condition” (403). In these projects, the respondents were divided into two groups, one group assessed using the ATA Framework and the other assessed using the LBI Bandscale. Because of this design, the t-test used is the independent samples t-test. The independent samples t-test is used to analyze a two-sample experiment consisting of independent samples (Ibid.). Below is the table showing the results of the t-test of the pilot project in Fall 2013.

Table 6.5
T-Test Results of the Fall 2013 Pilot Project

<table>
<thead>
<tr>
<th>Group Statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
</tr>
<tr>
<td>----</td>
</tr>
<tr>
<td>7</td>
</tr>
<tr>
<td>5</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Independent Samples Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Levene's Test for Equality of Means</td>
</tr>
<tr>
<td>F</td>
</tr>
<tr>
<td>-----</td>
</tr>
<tr>
<td>EVNA</td>
</tr>
</tbody>
</table>

Notes: EVA = Equal Variances Assumed; EVNA = Equal Variances Not Assumed

The table above indicates that the total number of the respondents (N) is 12 (7+5), but it is not N which determines the appropriate t-distribution for a study (Heiman, 2001:375). It is df or degrees of freedom that determine it. In the two-tailed test and for this independent samples t-test, the df is N – 2, so the df is 10 in this project. A t-distribution is “a theoretical sampling distribution of all possible values of $t_{obs}$ when a raw score population is infinitely sampled using a particular N” (393). A two-tailed test is a test used “when the direction in
which the dependent scores will change is not predicted” (363). The result obtained ($t_{obt}$) is
$t(10) = .087$, $p>.05$. Referring to the table of Critical Values of $t$: The $t$-Tables from E. Pearson and H. Hartley, the $t$ critical value ($t_{crit}$) with $df$ 10 and $p>.05$ (95% confidence) is
$\pm2.228$ (taken from Heiman, 2001:708), so to be able to reject the null hypothesis the $t_{obt}$ should be beyond $t_{crit}$. The null hypothesis ($H_0$) is “the statistical hypothesis that describes the population $\mu_s$ (the population mean) being represented if the predicted relationship does not exist” (363). The result shows there is no significant difference between the average of ED based on the ATA Framework and the average of ED based on the LBI Bandscale because the $t_{obt}$ (0.87) does not lie in the region of rejection which is beyond the $t_{crit}$ ($\pm2.228$). This also means the null hypothesis for this calculation cannot be rejected. In other words, according to the result of this pilot project of Fall 2013, the two approaches do not display a significant difference in their effectiveness. The graph below shows the sampling distribution for this result.

As for the results of the Spring 2014, there are only data from the respondents who were assessed using the ATA Framework. Out of three who did the first assignment, two returned the final assignment. In contrast, only two who did the first assignment were assessed using the LBI Bandscale, and none returned their revisions (the final assignment).
This situation makes it impossible to compare the ED based on both assessment models. Consequently, the data from Spring 2014 cannot be calculated and analyzed, so there is no discussion about the results from this term.

6.2.2 Weaknesses of the Pilot Projects and Changes for the Formal or Second Projects

The pilot projects revealed problems and obstacles which allowed for improving the design of the second projects. The weaknesses of the pilot projects are as follows. First, the design of the pilot projects is the between-subjects design where there are two groups of respondents, one group assessed using the ATA Framework and the other assessed using the LBI Bandscale. Those two groups only experienced one type of the assessment model applied in the study. In other words, the respondents did not see both assessment models. This condition resulted in the respondents believing that both assessment models are effective, as this can be seen in the discussion of the online survey results below. The problem is that the respondents could not decide which one of the two assessment models is more effective if they only saw one type. Furthermore, the LBI translation class students do not usually receive written feedback on their translations since all feedback and discussion about their translations are conducted in class. Therefore, the respondents were very excited and happy to receive the feedback in this research, either using the ATA Framework or using the LBI Bandscale.

The second weakness of the pilot projects is the bias that might occur due to the fact that I was the only rater or the one who assessed the translations and the revisions. It is true that I had an English native speaker (my advisor)’s supervision in assessing the translations, but officially there was only one rater in the pilot projects. As a result, another rater (an independent judge) was required to minimize any bias. The third weakness with this design is
that the data set obtained was small, and there were an imbalanced number of respondents between those who were assessed using the ATA Framework and those assessed using the LBI Bandscale, as seen in the Fall 2013 and the Spring 2014 data (see Table 6.4).

Therefore, to overcome the problems described above, there were some changes required to improve the design of the research for the formal projects. First of all, in order to make it possible for the respondents to see both assessment models, they were asked to translate two paragraphs; one paragraph was assessed using one type of the assessment model, and the other was assessed using the other type. This design is called the within-subjects design. A within-subjects design is a design where “for each participant in one condition there is a comparable participant in the other conditions” (Heiman, 2001:438). This first change automatically solved the third problem, as there were more data generated from this design. The second change was to have two raters for the revisions or the final assessments. Ideally two raters should have been involved starting with the translation (the first assignment) phase; however, this is too much to ask from the independent judge who lives and works in Jakarta, Indonesia. Moreover, having two raters for the first assignment (the translation) would probably confuse the respondents in doing their revisions. Thus, the second rater was only involved in the final assignment phase.

6.3 Formal or Second Project Result Analysis and Discussion

It is important to first discover the inter-rater reliability between me (Rater 1) and the independent judge (Rater 2) in assessing the revisions before the discussion of the second project data analysis. This correlation reveals how effective and how valid the first rater (me) was in evaluating all the assignments in all the projects. If the results show good inter-rater
reliability, it is possible to claim confidence for the validity and reliability of the results, as well as for the data analysis and interpretation.

6.3.1 Inter-Rater Reliability

As described above and in Chapter 5, in the formal projects an independent judge (the second rater) assessed the final assignments (the revisions). Thus, it is necessary to discuss the inter-rater reliability between Rater 1 (me) and Rater 2 (the independent judge). Inter-rater reliability in statistics (also called inter-rater agreement or concordance) is the degree of agreement among raters, and it provides a score to show the level of homogeneity or consensus in the ratings (Gwet, 2014:4-6). According to Michaelsen et al., “inter-rater reliability is related to the stability of the data obtained by two or more evaluators from the same measurement” (2011:81). The term inter-rater reliability should be applied in the measurement obtained through observation of performance where there is no physical contact between the raters and the respondents in the research (Ibid.). As this study is an observational study, the use of inter-rater reliability in the second projects is appropriate.

Inter-rater reliability is used to help create a degree of objectivity among subjective assessments made by several raters (http://study.com/academy/lesson/inter-rater-reliability-in-psychology-definition-formula-quiz.html). There are many statistical approaches that can be used to determine inter-rater reliability, such as Cohen’s kappa, Fleiss’ kappa, inter-rater correlation, concordance correlation coefficient, intra-class correlation, Pearson’s r Correlation, and others. For this study, I refer to Pearson’s r Correlation to discover whether there is a positive correlation between Rater 1 and Rater 2. Below is the table of Pearson’s r Correlation.
**Table 6.6**
Pearson’s \( r \) Correlation

<table>
<thead>
<tr>
<th>( r )</th>
<th>Correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>+.70 or higher</td>
<td>Very strong positive relationship</td>
</tr>
<tr>
<td>+.40 to +.69</td>
<td>Strong positive relationship</td>
</tr>
<tr>
<td>+.30 to +.39</td>
<td>Moderate positive relationship</td>
</tr>
<tr>
<td>+.20 to +.29</td>
<td>Weak positive relationship</td>
</tr>
<tr>
<td>+.01 to +.19</td>
<td>No or negligible relationship</td>
</tr>
<tr>
<td>-.01 to -.19</td>
<td>No or negligible relationship</td>
</tr>
<tr>
<td>-.20 to -.29</td>
<td>Weak negative relationship</td>
</tr>
<tr>
<td>-.30 to -.39</td>
<td>Moderate negative relationship</td>
</tr>
<tr>
<td>-.40 to -.69</td>
<td>Strong negative relationship</td>
</tr>
<tr>
<td>-.70 or lower</td>
<td>Very strong negative relationship</td>
</tr>
</tbody>
</table>

Source: [http://faculty.quinnipiac.edu/libarts/polsci/Statistics.html](http://faculty.quinnipiac.edu/libarts/polsci/Statistics.html)

Based on the table above, the following explains the significance of the correlation with several examples. The positive relationship or correlation means there is homogeneity or a similarity in assessing or evaluating something. For instance, when there is a positive correlation between two raters in assessing translations, it means both agree with all or almost all of what is considered “successful or functional” and/or “not successful or not functional” in the translations. No or negligible relationship refers to no correlation at all. For example, when the two raters have no correlations, it means they do not agree on any standard as they do not have the same ideas on what is “successful” or “not successful” in assessing something. Furthermore, the negative relationship or correlation means the relationship is inverse; for example, when Rater 1 gives A (excellent grade), Rater 2 will give D (very bad grade), or the other way around. Although there is a correlation, this shows that there is no significant validity between the raters. Thus, a positive correlation is desirable when evaluating inter-rater reliability.
In order to discover the correlation of the inter-rater reliability between Rater 1 and Rater 2 in the formal projects, several calculations were performed based on the table below consisting of the data from Summer 2014, Fall 2014, and Spring 2015.

Table 6.7  
Second Project Data

<table>
<thead>
<tr>
<th>Respondent</th>
<th>ATA Framework</th>
<th>LBI Bandscale</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>TE</td>
<td>RE1</td>
</tr>
<tr>
<td>Sum12014</td>
<td>48</td>
<td>26</td>
</tr>
<tr>
<td>Sum22014</td>
<td>24</td>
<td>7</td>
</tr>
<tr>
<td>Sum32014</td>
<td>23</td>
<td>12</td>
</tr>
<tr>
<td>Sum42014</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>Fall12014</td>
<td>43</td>
<td>27</td>
</tr>
<tr>
<td>Fall22014</td>
<td>21</td>
<td>10</td>
</tr>
<tr>
<td>Fall32014</td>
<td>34</td>
<td>22</td>
</tr>
<tr>
<td>Fall42014</td>
<td>9</td>
<td>2</td>
</tr>
<tr>
<td>Fall52014</td>
<td>28</td>
<td>17</td>
</tr>
<tr>
<td>Fall62014</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>Spr12015</td>
<td>15</td>
<td>13</td>
</tr>
<tr>
<td>Spr22015</td>
<td>16</td>
<td>4</td>
</tr>
<tr>
<td>Spr32015</td>
<td>23</td>
<td>10</td>
</tr>
<tr>
<td>Mean</td>
<td>23.62</td>
<td>13.15</td>
</tr>
</tbody>
</table>

Notes: TE = Translation Errors; RE = Revision Errors; ED = Error Difference

Based on the table above, ED1 is the error difference obtained from the number of translation errors subtracted from the number of revision errors 1 (RE1). RE1 was obtained as the result of Rater 1’s assessment.

\[ \text{TE} - \text{RE1} = \text{ED1} \]

ED2 is the error difference obtained from the number of translation subtracted from the number of revision errors 2 (RE2). RE2 was the result of Rater 2’s assessment.

\[ \text{TE} - \text{RE2} = \text{ED2} \]
There are three correlations obtained from the above data. First, all ED1s for the ATA Framework and the LBI Bandscale are compared with all ED2s for the ATA Framework and the LBI Bandscale. The result can be seen in the table below.

**Table 6.8**  
Correlation of All ED1s and ED2s for the ATA Framework and the LBI Bandscale

<table>
<thead>
<tr>
<th></th>
<th>ED1</th>
<th>ED2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pearson Correlation</td>
<td>1.00</td>
<td>.754**</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>.000</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>26.00</td>
<td>26.00</td>
</tr>
<tr>
<td>Pearson Correlation</td>
<td>.754**</td>
<td>1.00</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>.000</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>26.00</td>
<td>26.00</td>
</tr>
</tbody>
</table>

**. Correlation is significant at the 0.01 level (2-tailed).

The table shows that \( r(24) = .754, p < .01 \), and referring to Table 6.6 above for the Pearson \( r \) Correlation, this result shows a very strong positive relationship or correlation. It means Rater 1 and Rater 2 have the same ideas about what is ‘successful or functional’ and ‘not successful or not functional’ in translations, or at least the number of errors found in translations by those two raters is mostly similar. In addition, this shows good inter-rater reliability. The assessment of both Rater 1 and Rater 2 can be considered valid and reliable, at least in the value of ED.

The second correlation is between ED1 and ED2 for the ATA Framework. The result can be seen in the following table.

108
Table 6.9
Correlation of ED1 and ED2 for the ATA Framework

<table>
<thead>
<tr>
<th></th>
<th>ED1</th>
<th>ED2</th>
</tr>
</thead>
<tbody>
<tr>
<td>ED1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pearson Correlation</td>
<td>1.00</td>
<td>.841**</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>.000</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>13.00</td>
<td>13.00</td>
</tr>
<tr>
<td>ED2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pearson Correlation</td>
<td>.841**</td>
<td>1.00</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>.000</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>13.00</td>
<td>13.00</td>
</tr>
</tbody>
</table>

**. Correlation is significant at the 0.01 level (2-tailed).

Based on the table above, the result is \( r(11) = .841, p < .01 \), and this also shows a very strong positive correlation. The third correlation is between ED1 and ED2 for the LBI Bandscale, and the result can be seen below.

Table 6.10
Correlation of ED1 and ED2 for the LBI Bandscale

<table>
<thead>
<tr>
<th></th>
<th>ED1</th>
<th>ED2</th>
</tr>
</thead>
<tbody>
<tr>
<td>ED1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pearson Correlation</td>
<td>1.00</td>
<td>.689**</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>.009</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>13.00</td>
<td>13.00</td>
</tr>
<tr>
<td>ED2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pearson Correlation</td>
<td>.689**</td>
<td>1.00</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>.009</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>13.00</td>
<td>13.00</td>
</tr>
</tbody>
</table>

The table above shows that the result is \( r(11) = .689, p < .01 \), and based on Pearson \( r \) Correlation this means a strong positive relationship. Even though the \( r \) of the ED1 and ED2 for the LBI Bandscale is somewhat lower than the \( r \) of the ED1 and ED2 for the ATA Framework and than the \( r \) of all ED1s and ED2s for both the ATA Framework and the LBI Bandscale, it still shows that there is a strong positive correlation. These results confirm the reliability of Rater 1, so it is unnecessary to include Rater 2 in order to assess the reliability of the assessments. In spite of that, the data analyses below will account for both raters.
6.3.2 The Quantitative Analysis and the Interpretation of the Second Projects

Based on Table 6.7 showing the Second Project Data, there are several statistical operations that can be performed to show whether there is a significant difference between the improvement of translations (shown by ED – error difference) assessed using the ATA Framework and the ones assessed using the LBI Bandscale. The first calculation is based on all ED1s for the ATA Framework and the LBI Bandscale. Referring to Table 6.7, the mean of ED 1 for the ATA Framework is 10.46 or 47.54%. The mean of ED1 for the LBI Bandscale is 7.46 or 33.60%. This shows that improvement using the ATA Framework is higher than improvement using the LBI Bandscale. The different percentage of ED1 for both assessment models looks ostensibly as if the difference might be significant.

However, in the calculation using SPSS, the result is different. Because the design for the second projects is the within-subjects design, the t-test applied is the dependent samples t-test or the paired samples t-test. The dependent-samples t-test is “the parametric inferential procedure used to test either a matched-groups or repeated-measures design” (Heiman, 2001:449). A matched-groups design and a repeated-measures design are the same as the within-subjects design (438). In a matched-groups design, “each participant in one condition ‘matches’ a participant in the other condition(s) on one or more extraneous variable” (431), while in a repeated-measures design “each participant is tested under all conditions of an independent variable” (433). Based on the paired samples t-test for ED1 for both assessment models as seen in Table 6.11 below, the result of the \( t_{\text{obt}} \) is \( t(12) = 1.183, p>.05 \). The \( t_{\text{crit}} \) is \( \pm 2.179 \) for \( df \ 12 \) and \( p>.05 \) (Heiman, 2001:708). This means that the difference is not significant because the \( t_{\text{obt}} \) does not lie in the region of rejection or is not beyond the \( t_{\text{crit}} \). and
it also means that the null hypothesis cannot be rejected for this calculation. Figure 6.5 below shows the sampling distribution for this result.

### Table 6.11
The Paired Samples $T$-Test for ED1 for the ATA Framework and the LBI Bandscale

<table>
<thead>
<tr>
<th>Paired Samples Statistics</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>N</td>
<td>Std. Deviation</td>
<td>Std. Error Mean</td>
</tr>
<tr>
<td>Pair 1</td>
<td>ATA</td>
<td>10.4615</td>
<td>13</td>
<td>6.34580</td>
</tr>
<tr>
<td></td>
<td>LBI</td>
<td>7.4615</td>
<td>13</td>
<td>4.78914</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Paired Samples Correlations</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>Correlation</td>
</tr>
<tr>
<td>Pair 1</td>
<td>13</td>
<td>-.337</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Paired Samples Test</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Std. Deviation</td>
<td>Std. Error Mean</td>
<td>95% Confidence Interval of the Difference</td>
<td>t</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Lower</td>
<td>Upper</td>
</tr>
</tbody>
</table>

**Figure 6.5: Sampling Distribution of the ED1 of the Second Projects**

The second calculation examines all the ED2s for the ATA Framework and the LBI Bandscale. From Table 6.7, the mean of ED2 for the ATA Framework is 11.46 or 46.34%. 

111
while the mean of ED2 for the LBI Bandscale is 10.46 or 46.57%. From these means, there is only a very small difference between improvement with the ATA Framework and the one with the LBI Bandscale. In addition, based on the result of the $t$-test in the table below, it shows that the $t_{\text{obt}}$ is $t(12) = 0.279$, $p > 0.05$, whereas the $t_{\text{crit}}$ is $\pm 2.179$. This means that there is no significant difference in improvement based on the two assessment models. Figure 6.6 shows the sampling distribution for this result.

Table 6.12  
The Paired Samples $T$-Test for ED2 for the ATA Framework and the LBI Bandscale

<table>
<thead>
<tr>
<th>Paired Samples Statistics</th>
<th>Mean</th>
<th>N</th>
<th>Std. Deviation</th>
<th>Std. Error Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pair 1 ATA</td>
<td>11.4615</td>
<td>13</td>
<td>8.10982</td>
<td>2.24926</td>
</tr>
<tr>
<td>LBI</td>
<td>10.4615</td>
<td>13</td>
<td>8.26252</td>
<td>2.29161</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Paired Samples Correlations</th>
<th>N</th>
<th>Correlation</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pair 1 ATA &amp; LBI</td>
<td>13</td>
<td>-.247</td>
<td>.416</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Paired Samples Test</th>
<th>Mean</th>
<th>Std. Deviation</th>
<th>Std. Error Mean</th>
<th>95% Confidence Interval of the Difference</th>
<th>t</th>
<th>df</th>
<th>Sig. (2-tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pair 1 ATA - LBI</td>
<td>1.00000</td>
<td>12.92929</td>
<td>3.58594</td>
<td>-6.81309 - 8.81309</td>
<td>.279</td>
<td>12</td>
<td>.785</td>
</tr>
</tbody>
</table>
Figure 6.6: Sampling Distribution of the ED2 of the Second Projects

The following statistical calculations compare all paragraphs 1 assessed using the ATA Framework versus all paragraphs 1 assessed using the LBI Bandscale and all paragraphs 2 (the ATA Framework versus the LBI Bandscale). Since each paragraph was only assessed using one assessment model, the $t$-test applied was the independent samples $t$-test because the design is the between-subjects design in these calculations. The third calculation compares ED1 for paragraph 1 assessed using the ATA Framework to ED1 for paragraph 1 assessed using the LBI Bandscale. The fourth calculation reflects the ED2 of paragraph 1 assessed using both assessment models. The tables below will show the data for all paragraphs 1 and the $t$-test results.

<table>
<thead>
<tr>
<th>Table 6.13</th>
<th>Paragraph 1 Data</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Second Project Results - Paragraph 1 with the ATA Framework</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Respondents</strong></td>
<td>TE</td>
</tr>
<tr>
<td>1 Fall12014</td>
<td>43</td>
</tr>
<tr>
<td>1 Fall32014</td>
<td>34</td>
</tr>
<tr>
<td>1 Fall52014</td>
<td>28</td>
</tr>
<tr>
<td>1 Spr22015</td>
<td>16</td>
</tr>
<tr>
<td><strong>Average/Mean</strong></td>
<td>30.29</td>
</tr>
</tbody>
</table>
### Second Project Results - Paragraph 1 with the LBI Bandscale

<table>
<thead>
<tr>
<th>Respondents</th>
<th>TE</th>
<th>RE1</th>
<th>ED1</th>
<th>% ED1</th>
<th>RE2</th>
<th>ED2</th>
<th>% ED2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Sum22014</td>
<td>34</td>
<td>23</td>
<td>11</td>
<td>32.35</td>
<td>15</td>
<td>19</td>
<td>55.88</td>
</tr>
<tr>
<td>1 Fall2014</td>
<td>34</td>
<td>20</td>
<td>14</td>
<td>41.18</td>
<td>6</td>
<td>28</td>
<td>82.35</td>
</tr>
<tr>
<td>1 Fall42014</td>
<td>24</td>
<td>14</td>
<td>10</td>
<td>41.67</td>
<td>7</td>
<td>17</td>
<td>70.83</td>
</tr>
<tr>
<td>1 Fall62014</td>
<td>15</td>
<td>12</td>
<td>3</td>
<td>20.00</td>
<td>2</td>
<td>13</td>
<td>86.67</td>
</tr>
<tr>
<td>1 Spr12014</td>
<td>29</td>
<td>18</td>
<td>11</td>
<td>37.93</td>
<td>18</td>
<td>11</td>
<td>73.93</td>
</tr>
<tr>
<td>1 Spr32014</td>
<td>27</td>
<td>26</td>
<td>1</td>
<td>3.70</td>
<td>21</td>
<td>6</td>
<td>22.22</td>
</tr>
<tr>
<td><strong>Average/Mean</strong></td>
<td>27.17</td>
<td>18.83</td>
<td>8.33</td>
<td>29.47</td>
<td>11.50</td>
<td>15.67</td>
<td>59.31</td>
</tr>
</tbody>
</table>

Notes: TE = Translation Errors; RE = Revision Errors; ED = Error Difference

Based on Table 6.13, the mean for paragraph 1 ED1 for the ATA Framework is 12 or 40.06%, while the mean of paragraph 1 ED1 of the LBI Bandscale is 8.33 or 29.47%. These data would indicate that there is a difference in effectiveness and that the ATA Framework seems to be more effective than the LBI Bandscale. However, from the $t$-test in Table 6.14 below, the result of the $t_{obt}$ is $t(11) = 1.102$, $p > 0.05$, while the $t_{crit}$ is ±2.201 (Heiman, 2001:708). This indicates that the difference is not significant. Figure 6.7 shows the sampling distribution for this result.

#### Table 6.14

**The Independent Samples $T$-Test for Paragraph 1 ED1 for the ATA Framework and the LBI Bandscale**

<table>
<thead>
<tr>
<th>Group Statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>VAR00002</td>
</tr>
<tr>
<td>ATA</td>
</tr>
<tr>
<td>LBI</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Independent Samples Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Levene's Test for Equality of</td>
</tr>
<tr>
<td>Sig.</td>
</tr>
<tr>
<td>VAR00001</td>
</tr>
</tbody>
</table>

Notes: EVA = Equal Variances Assumed; EVNA = Equal Variances Not Assumed
Figure 6.7: Sampling Distribution of the Paragraph 1 ED1 of the Second Projects

For paragraph 1 ED2, based on Table 6.13, the mean of the ATA Framework is 13.14 or 41.84%, whereas the mean of the paragraph 1 ED2 of the LBI Bandscale is 15.67 or 59.31%. This shows that there is a difference between the ATA Framework and the LBI Bandscale, and that the LBI Bandscale appears to be more effective than the ATA Framework. Nevertheless, based on Table 6.15 below, the result of the $t_{obt}$ is $t(11) = -0.540$, $p > 0.05$, while the $t_{crit}$ is $\pm 2.201$. Thus, this analysis of the data reveals that there is no significant difference between those two assessment models. Figure 6.8 shows the sampling distribution for this result.

Table 6.15
The Independent Samples $T$-Test for Paragraph 1 ED2 for the ATA Framework and the LBI Bandscale

<table>
<thead>
<tr>
<th>Group Statistics</th>
<th>VAR00002</th>
<th>N</th>
<th>Mean</th>
<th>Std. Deviation</th>
<th>Std. Error Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>VAR00001</td>
<td>1.00</td>
<td>7</td>
<td>13.1429</td>
<td>9.02642</td>
<td>3.41166</td>
</tr>
<tr>
<td>VAR00001</td>
<td>2.00</td>
<td>6</td>
<td>15.6667</td>
<td>7.58068</td>
<td>3.09480</td>
</tr>
</tbody>
</table>

115
Independent Samples Test

<table>
<thead>
<tr>
<th></th>
<th>Levene's Test for Equality of Variances</th>
<th>t-test for Equality of Means</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>F</td>
<td>Sig.</td>
</tr>
<tr>
<td>VAH00001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EVA</td>
<td>.253</td>
<td>.625</td>
</tr>
<tr>
<td>EVNA</td>
<td>-.548</td>
<td>11.000</td>
</tr>
</tbody>
</table>

Notes: EVA = Equal Variances Assumed; EVNA = Equal Variances Not Assumed

Figure 6.8: Sampling Distribution of the Paragraph 1 ED2 of the Second Projects

The fifth calculation deals with paragraph 2 ED1 for the ATA Framework and the LBI Bandscale. Based on Table 6.16 below, the mean of paragraph 2 ED1 for the ATA Framework is 8.67 or 56.25%. The mean of paragraph 2 ED1 for the LBI Bandscale is 6.71 or 37.13%. These values imply a large difference between the ATA Framework and the LBI Bandscale, and that the ATA Framework seems to be more effective than the LBI Bandscale. However, based on Table 21 below, the $t_{obt}$ is $t(11) = 0.650$, $p > 0.05$. This means that the difference is still not significant. Figure 6.9 below shows the sampling distribution for this result.
### Table 6.16
**Paragraph 2 Data**

<table>
<thead>
<tr>
<th>Respondents</th>
<th>TE</th>
<th>RE1</th>
<th>ED1</th>
<th>% ED1</th>
<th>RE2</th>
<th>ED2</th>
<th>% ED2</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 Sum2014</td>
<td>24</td>
<td>7</td>
<td>17</td>
<td>70.83</td>
<td>13</td>
<td>11</td>
<td>45.83</td>
</tr>
<tr>
<td>2 Fall2014</td>
<td>21</td>
<td>10</td>
<td>11</td>
<td>52.38</td>
<td>4</td>
<td>17</td>
<td>80.95</td>
</tr>
<tr>
<td>2 Fall42014</td>
<td>9</td>
<td>2</td>
<td>7</td>
<td>77.78</td>
<td>2</td>
<td>7</td>
<td>77.78</td>
</tr>
<tr>
<td>2 Fall62014</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td>66.67</td>
<td>3</td>
<td>0</td>
<td>0.00</td>
</tr>
<tr>
<td>2 Spr12014</td>
<td>15</td>
<td>13</td>
<td>2</td>
<td>13.33</td>
<td>11</td>
<td>4</td>
<td>26.67</td>
</tr>
<tr>
<td>2 Spr32014</td>
<td>23</td>
<td>10</td>
<td>13</td>
<td>56.52</td>
<td>5</td>
<td>18</td>
<td>78.26</td>
</tr>
<tr>
<td><strong>Average/Mean</strong></td>
<td>15.83</td>
<td>7.17</td>
<td>8.67</td>
<td>56.25</td>
<td>6.33</td>
<td>9.50</td>
<td>51.58</td>
</tr>
</tbody>
</table>

### Table 6.17
**The Independent Samples T-Test for Paragraph 2 ED1 for the ATA Framework and the LBI Bandscale**

<table>
<thead>
<tr>
<th>Group Statistics</th>
<th>VAR00002</th>
<th>N</th>
<th>Mean</th>
<th>Std. Deviation</th>
<th>Std. Error Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>VAR00001</td>
<td>1.00</td>
<td>6</td>
<td>8.6667</td>
<td>6.08824</td>
<td>2.48551</td>
</tr>
<tr>
<td></td>
<td>2.00</td>
<td>7</td>
<td>6.7143</td>
<td>4.75094</td>
<td>1.79569</td>
</tr>
</tbody>
</table>

**Independent Samples Test**

<table>
<thead>
<tr>
<th>Levene's Test for Equality of Variances</th>
<th>t-test for Equality of Means</th>
</tr>
</thead>
<tbody>
<tr>
<td>F</td>
<td>Sig.</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>VAR00001</td>
<td>EVA</td>
</tr>
<tr>
<td></td>
<td>EVNA</td>
</tr>
</tbody>
</table>

Notes: EVA = Equal Variances Assumed; EVNA = Equal Variances Not Assumed
The sixth calculation shows paragraph 2 ED2 for both assessment models. Table 6.16 above shows that the mean of paragraph 2 ED2 for the ATA Framework is 9.50 or 51.58%. The mean of paragraph 2 ED2 for the LBI Bandscale is 6.00 or 35.65%. This demonstrates that there is a difference between the ATA Framework and the LBI Bandscale, and that the ATA Framework appears to be more effective than the LBI Bandscale. Nonetheless, based on Table 6.18 below, the result of the \( t_{obt} \) is \( t(11) = 0.945, p > 0.05 \). This means there is no significant difference between the ATA Framework and the LBI Bandscale. Consequently, the null hypothesis cannot be rejected. Figure 6.10 below shows the sampling distribution for this result.

**Table 6.18**
The Independent Samples \( T \)-Test for Paragraph 2 ED2 for the ATA Framework and the LBI Bandscale

<table>
<thead>
<tr>
<th>Group Statistics</th>
<th>VAR00002</th>
<th>N</th>
<th>Mean</th>
<th>Std. Deviation</th>
<th>Std. Error Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>VAR000002</td>
<td>1.00</td>
<td>6</td>
<td>9.500</td>
<td>7.17635</td>
<td>2.92973</td>
</tr>
<tr>
<td>VAR000001</td>
<td>2.00</td>
<td>7</td>
<td>6.000</td>
<td>6.19139</td>
<td>2.34013</td>
</tr>
</tbody>
</table>

118
Independent Samples Test

<table>
<thead>
<tr>
<th></th>
<th>F</th>
<th>Sig.</th>
<th>t</th>
<th>df</th>
<th>Sig. (2-tailed)</th>
<th>Mean Difference</th>
<th>Std. Error Difference</th>
<th>95% Confidence Interval of the Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>VAH00001</td>
<td></td>
<td></td>
<td>EVA</td>
<td></td>
<td>.413</td>
<td>.365</td>
<td>3.50000</td>
<td>-4.65183 to 11.65183</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>EVNA</td>
<td></td>
<td>.933</td>
<td>.373</td>
<td>3.50000</td>
<td>-4.85268 to 11.85268</td>
</tr>
</tbody>
</table>

Notes: EVA = Equal Variances Assumed; EVNA = Equal Variances Not Assumed

Figure 6.10: Sampling Distribution of the Paragraph 2 ED2 of the Second Projects

Based on Tables 6.13 and 6.16, almost all respondents made improvement in their translations or had fewer errors in the final assignments or the revisions. However, there are a few outliers. An outlier is “an observation that lies an abnormal distance from other values in a random sample from a population” (http://www.itl.nist.gov/div898/handbook/prc/section1/prc16.htm). An outlier point does not follow the general characteristics or traits of other points. In these projects, outliers are those points that do not show any improvement or any decreased errors. Instead, they show zero improvement and an increase in errors. Zero improvement means 0%, as there is no increase or decrease in errors. An increase in errors means there are more errors in the RE than in the TE, and in the calculation the value for ED falls in the minus range, which represents an increase in errors.
For paragraph 1 (Table 6.13), there is one outlier from the respondent Sum42014, and this paragraph was assessed using the ATA Framework. The ED1 of this respondent is 0 or no improvement, and the ED2 is -2 or -10%, which means there is an increased number of errors. For paragraph 2 (Table 6.16), there are three outliers. First is the respondent Fall62014, and this paragraph was assessed using the ATA Framework. Despite having an error decrease based on the ED1, the ED2 of this respondent is 0. Second is the ED2 of the respondent Sum12014 which shows an increase in errors, as the number is -3 or -12.5%, although the ED1 shows a decrease in errors with 4 points or 16.67%. This paragraph was assessed using the LBI Bandscale. The last outlier is from the respondent Fall32013, as the ED1 shows no improvement or 0%, while the ED2 reveals an decrease in errors with 2 points or 13.33%. The paragraph was assessed using the LBI Bandscale.

It is not known why there are several outliers. The reason(s) for these outliers could be revealed if the researcher had been the teacher in class, as a teacher can have direct communication and provide direct feedback to the students. The two assessment models have nothing to do with these outlier results because we can see above that two outlier paragraphs (1 and 2) were assessed using the ATA Framework, and the other two (all paragraphs 2) were assessed using the LBI Bandscale. Nevertheless, it is interesting to discover that only one outlier is found in paragraph 1, whereas there are three outliers in paragraph 2. This might show that paragraph 2 is potentially more difficult than paragraph 1.

Besides the outliers, it is necessary to discuss why there is no calculation that compares all paragraphs 1 versus all paragraphs 2. The reason is because the length of paragraph 1 is not the same as the length of paragraph 2. As described in Chapter 5, paragraph 1 consists of 183 words, while paragraph 2 consists of 131 words. Hence, paragraph 1 is longer than paragraph 2 by 52 words. To compare them will not be valid. The potential for even a small difference in difficulty between the two paragraphs is another
reason not to try to come up with any statistical comparison between the two. These two paragraphs, however, are part of the same text, so the language and the difficulty of those paragraphs should be similar.

Although the results of the means (or the average) based on the data of the pilot and formal projects mostly show that the ATA Framework has more effectiveness than the LBI Bandscale, all the statistical calculations with $t$-test always result in not significant differences of the effectiveness between the two tools. This reveals that despite of the fact that there is a difference, the difference is not large. This might occur due to the small number of the respondents in this study. If the number of the respondents had reached above 100, for instance, the results of the data could potentially lead to significant differences between the two tools. Hence, further research is required to acquire more respondents and enlist more data. The following subchapter will discuss the online survey results.

**6.4 Online Survey Results**

As described in Chapter 5, the survey was conducted online using Survey Monkey. The survey for the pilot projects is different from the survey for the second projects because of the design change. Nonetheless, the purpose of both surveys is the same, which is to discover the perspectives of the respondents, who are translation students, on translation assessment, especially the two models, the ATA Framework and the LBI Bandscale. This subchapter describes the results of the two surveys.
6.4.1 Pilot Project Survey Results

The title of the pilot project survey is “Perspectives on Translation Assessment.” There were 11 respondents filling in this survey. There are only six questions in this survey. The respondents had already been asked to do several things in this study, so the survey was designed to be short and simple so that the respondents would not feel burdened with too many tasks to perform. The details of the pilot project online survey can be seen in Appendix E.

The first question asks what type of assessment the respondents received. There are two choices given: A. group 1 (the ATA Framework) and B. group 2 (the LBI Bandscale). Out of 11 participants, seven people belonged to group 1, and four belonged to group 2. The second question asks how general or specific the assessment given to the respondents was. There are four answers provided: A. very general, B. general, C. specific, and D. very specific. “General” here refers to having no detailed explanation, and “specific” means having detailed explanation. Below is the graph describing the answers of this question.

**Figure 6.11**
How General or Specific the Assessment Models
Based on the graph above, one respondent chose “very general,” two chose “general,” four chose “specific,” and four chose “very specific.” However, the graph above does not really show which assessment type is considered general or specific. The pie charts below will reveal what respondents think of each type of assessment. The first pie chart is for the ATA Framework, and the second pie chart is for the LBI Bandscale.

![Pie Chart ATA Framework](image)

**Figure 6.12**

The ATA Framework: General or Specific?

![Pie Chart LBI Bandscale](image)

The LBI Bandscale: General or Specific?

Based on the first pie chart above, no respondent thinks that the ATA Framework is very general, although there is one person considering this model as general. Three respondents think that this model is specific, and the other three consider that this model is
very specific. Overall, most respondents believe the ATA Framework is specific. For the second pie chart, it is rather confusing because each respondent in this group 2 chose a different answer. One person chose “very general,” another chose “general,” another chose “specific,” and the last person chose “very specific.” Consequently, the result is inconclusive for the LBI Bandscale.

The third question asks whether the assessment models are simple or complicated. The terms “simple” and “complicated” are relative in their meaning. In this survey, the meaning of these terms is based on the understanding of the respondents who are students, and the meaning might be different from the points of view of the teachers and the raters who use the two assessment models. The term “simple” here could mean that the assessment model is easy for the respondents to understand and use to improve their translations, while the term “complicated” might mean that the assessment model is difficult for the respondents to understand and use. The pie charts below will show the results of the third question. The first pie chart illustrates data for the ATA Framework. The second pie chart refers to the LBI Bandscale.

Figure 6.13

The ATA Framework: Simple or Complicated?
The first pie chart shows that five respondents think that the ATA Framework is simple. One person chose “complicated,” and another one chose “very complicated.” Thus, in general, the ATA Framework is considered simple by most of the respondents. The second pie chart shows that two persons think the LBI Bandscale is very simple and the other two believe it is simple. The conclusion is that the LBI Bandscale is simple. Since the respondents assessed with the ATA Framework did not see the LBI Bandscale, and the other way around, most of the two groups of respondents think both assessment models are simple or have no differences.

Questions 4, 5, and 6 are related to one another. Question 4 asks whether the assessment model provided to the respondents is helpful or not in improving their translations. If the respondents answer that it is useful, then they just need to give the reason(s) in number 5 and skip number 6. If they think it is not useful, then they should provide the reason(s) in number 6 and skip number 5. In this pilot project survey, all respondents believe that the two assessment models are helpful in improving their translations, so every respondent answered number 4 with yes, and 10 out of 11 provided their reasons in number 5. One person did not give his/her reason. Below is the table showing their reasons and they are recorded as they are without any changes.
Table 6.19
Reasons Why the Assessment Models are Useful

<table>
<thead>
<tr>
<th>The ATA Framework</th>
<th>The LBI Bandscale</th>
</tr>
</thead>
<tbody>
<tr>
<td>• The feed I received was that I needed to improve my use of tenses and articles, which was a good point. I'll definitely need to brush up those points.</td>
<td>• The assessment explanation given was very clear, both of the grammar, punctuation, usage, and the syntax. It is very helpful for me to know what grammar should be used to translate the news, besides it also add my knowledge about how to use the punctuation properly so that there was no misunderstanding appear.</td>
</tr>
<tr>
<td>• I understand that I am lack in reading English materials. That’s why I need to improve my English skill.</td>
<td>• I’ve had some difficulties with the assessment, fortunately your feedback has been really helpful to point my mistakes.</td>
</tr>
<tr>
<td>• To measure my translation skills to the details, and I found that the assessment was simple but need more attention and practice.</td>
<td>• To think of better translation for a particular sentence and to use them as a reference to not to repeat the same mistake again in the future when translating a document or article.</td>
</tr>
<tr>
<td>• The assessment and feedback I received was clear and very detail, it explained each mistakes I made and had enlighten me on how much I need to work more on my translation skill to achieve the required level for a good translator.</td>
<td>• I know which part of my translation to be revised. Thank you.</td>
</tr>
<tr>
<td>• Help me find the mistakes I made, be more careful in using the words and using the correct form, pay more attention to the meaning after it is translated.</td>
<td>• I’ll be more cautious in doing the translation, and I’ll try my best to use the assessment and feedback as my “check point” to check the result of my translation.</td>
</tr>
<tr>
<td>• I’ll be more cautious in doing the translation, and I’ll try my best to use the assessment and feedback as my “check point” to check the result of my translation.</td>
<td></td>
</tr>
</tbody>
</table>

From the table above, we can see that all respondents provided good reasons and perspectives on the two assessment models. They think both models can improve their translation skills, particularly in the target language (English) writing. They all show a positive attitude and appreciation towards the assessment given to them, and there is no negativity at all. Almost all the results from this survey show similar perspectives of the respondents towards both assessment models. This, however, might happen due to the fact that each respondent only saw and experienced one type of assessment.
6.4.2 Formal or Second Project Survey Results

The second project online survey’s title is “Comparing Two Translation Assessment Models.” It also used Survey Monkey. There were 11 respondents filling in this survey. It has seven questions: two closed-ended questions with two answers provided each, four questions using a Likert scale with five answers provided, and one open-ended question. The details of the survey can be seen in Appendix E. The first two questions ask what type of assessment was used in the paragraphs translated. The first one is about paragraph 1 and the second one is about paragraph 2. The graph below shows the results from these two questions.

![Figure 6.14 The Results of Questions 1 and 2](image)

The graph above shows that seven paragraphs 1 were assessed using the ATA Framework, and four paragraphs 1 were assessed using the LBI Bandscale. In contrast, four paragraphs 2 were assessed using the ATA Framework, and seven paragraphs 2 were assessed using the LBI Bandscale.

Questions 3 and 4 are exactly the same questions. They ask how difficult it is for the respondents to understand and apply the assessment. Question 3 refers to the ATA Framework, and question 4 refers to the LBI Bandscale. For question 3, the results are two (18.18%) respondents chose “very simple,” four (36.36%) chose “simple,” four (36.36%)
chose “normal,” one (9.09%) chose “complicated,” and no one chose “very complicated.” It can be concluded that most respondents think the ATA Framework is between normal and simple. For question 4, the results are no respondent chose “very simple,” two (18.18%) chose “simple,” four (36.36%) chose “normal,” five (45.45%) chose “complicated,” and none chose “very complicated.” In general, most respondents think that the LBI Bandscale is between normal and complicated. Below is the graph describing the results of questions 3 and 4.

**Figure 6.15**
The Results of Questions 3 and 4

Questions 5 and 6 are also similar, but the difference is that question 5 is about the ATA Framework, while question 6 is about the LBI Bandscale. These questions ask how effective the assessment is in helping the respondents revise their translations. For question 5, the results indicate that no respondent chose “a great deal less effective,” also no one chose “somewhat less effective,” three (27.27%) respondents chose “effective,” four (36.36%) chose “somewhat more effective,” and four (36.36%) chose “a great deal more effective.” Hence, the respondents from the second projects consider that the ATA Framework is between “effective” and “a great deal more effective.” For question 6, the results are no
respondent chose “a great deal less effective,” five (45.45%) chose “somewhat less effective,” six (54.55%) chose “effective,” no one chose “somewhat more effective,” and none chose “a great deal more effective.” It is quite difficult to decide the results for the LBI Bandscale because almost half of the respondents think it is less effective (than the ATA Framework), while the other half believe it is effective. The graph below will describe the results of questions 5 and 6.

**Figure 6.16**
The Results of Questions 5 and 6

![](image)

Question 7 asks which method of assessment the respondents prefer for their translation assessment, the ATA Framework or the LBI Bandscale, and also asks for the reason(s). Ten out of 11 respondents prefer the ATA Framework to assess their translations. Only one prefers the LBI Bandscale. Nine respondents who prefer the ATA Framework wrote their reasons, but one person did not provide the reason. The person who chose the LBI Bandscale provided his/her reason(s). The table below contains the reasons written for this question. The reasons are recorded as they are without any changes.

129
### Table 6.20
The Results of Question 7

<table>
<thead>
<tr>
<th>Assessment Preference</th>
<th>Reason</th>
</tr>
</thead>
</table>
| The ATA Framework     | • ATA Framework because it gives detail explanation about my mistake so I can learn and revise it easily.  
                        | • ATA gives more detailed and precise assessment so it is easier to figure out and fix what goes wrong. LBI only points out the ones are wrong, but I gotta say it makes me rethink more creatively ways of delivering the message in the target language, but still leaves more room for error compared to ATA.  
                        | • ATA Framework despite its complicated details  
                        | • ATA Framework, because it explains detail things  
                        | • I prefer the ATA Framework because it is easier to find out what kind of translation errors that I have made, while by using the LBI Bandscale, I have to guess first what kind of errors that I made before revising it.  
                        | • The ATA Framework  
                        | • ATA Framework. It is more specific in pointing the error, makes it easier for me to decide the correct translation.  
                        | • The ATA Framework. It rightly pointed out the mistakes we made and clearly suggested the revision we should do.  
                        | • ATA Framework since it easy to be followed and easy to understand.  
                        | • I prefer to use ATA Framework as the assessment method because it’s more effective and make it easier to do the revision since the mistake point out clearly.  
                        | • The ATA Framework since it easy to be followed and easy to understand.  
                        | • I prefer to use ATA Framework as the assessment method because it’s more effective and make it easier to do the revision since the mistake point out clearly. |
| The LBI Bandscale     | • The LBI Bandscale because we also identify the mistake or miss of the its translation simpler the determine with the grade from A to D. |

Overall, this chapter has discussed the definitions of translation errors from several Translation Studies scholars as well as the definition used for these projects. In addition, the explanation of types of errors from a number of publications shows that there has been a variety of error categorizations, and this supports the formulation of the translation error typology used for these projects. The frequency of the error types from the projects in this study reveals from the most common error types to the least common ones in the Indonesian-English translations and revisions. In the analysis of the pilot project results and survey results, the interesting finding is that all respondents consider both the ATA Framework and the LBI Bandscale are effective and helpful in improving their translations. In the calculation of the mean of ED between the ATA Framework and the LBI Bandscale, the ATA
Framework seems to be slightly more effective, but based on the $t$-test the difference of the effectiveness between these two models is not significant.

In the analysis of the second project results and survey results, the important finding is the strong positive correlation between Rater 1 and Rater 2, so the inter-rater reliability is good in these projects. In the calculation of the means of all ED1 and ED2, most of the means show that there is a difference of effectiveness between the ATA Framework and the LBI Bandscale, and that the ATA Framework seems to be more effective than the LBI Bandscale. There is one mean showing that there is a very small difference of effectiveness between these two models, and there is one mean showing that the LBI Bandscale appears to be more effective than the ATA Framework. However, all the $t$-test results show that there is no significant difference between the two assessment models. In the survey results of these projects, most of the respondents prefer the ATA Framework, but they do not exactly consider the LBI Bandscale as ineffective. In the next chapter, the discussion will focus on the conclusions of this study related to the hypotheses, the limitations of the study, and the further research explained in the closing chapter.
CHAPTER 7

CONCLUSIONS:
SUMMARIES, LIMITATIONS, AND FUTURE RESEARCH

In general, based on the result analysis and discussion in Chapter 6, this study has demonstrated that both assessment models applied, the ATA Framework and the LBI Bandscale, are effective in helping students improve their translations. At first, because of the obvious differences between these two models, it was predicted there should have been a significant difference also in the effectiveness of the models. However, all the t-test results show there is no significant difference despite the fact that most of the means indicate apparent differences between the ATA Framework and the LBI Bandscale. This study also shows that there is a positive correlation between student perspectives and their translation results, and that the respondents in the formal projects prefer the ATA Framework to the LBI Bandscale, although they do not consider the LBI Bandscale ineffective.

7.1 Summaries

The summaries below will be related to the hypotheses proposed for this study. The first hypothesis is that there would be a significant difference in effectiveness between the ATA Framework and the LBI Bandscale. The second hypothesis is that students would prefer the ATA Framework to the LBI Bandscale in improving their translations, as indicated by the answers to the surveys. For the third hypothesis, it was predicted that there would be a
positive correlation between student perspectives on the assessment models and their translation results, based on the improvement made in the revisions and the comparison of survey results.

7.1.1 The Relative Effectiveness of the Two Models

As described above, based on the first hypothesis, this study attempts to discover the effectiveness of two different assessment models, the ATA Framework, an error analysis model, and the LBI Bandscale, a scale assessment model. It has also observed whether there is a significant difference in effectiveness between these two models. A summary of all the statistical calculations performed in Chapter 6 should reveal whether the first hypothesis is supported or rejected. The following table contains the summary of the statistical calculations.

<table>
<thead>
<tr>
<th>Name of the Project</th>
<th>Name of the Calculation</th>
<th>Mean in Percentage</th>
<th>$t_{obt}$</th>
<th>$t_{crit}$</th>
<th>Finding</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pilot Project</td>
<td>ED for ATA vs ED for LBI</td>
<td>61% vs 47.3%</td>
<td>$t(10) = .087$, p &gt; .05</td>
<td>±2.228</td>
<td>$H_0$ cannot be rejected.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Second Projects</td>
<td>All ED1 (ATA vs LBI)</td>
<td>47.54% vs 33.60%</td>
<td>$t(12) = 1.183$, p &gt; .05</td>
<td>±2.179</td>
<td>$H_0$ cannot be rejected.</td>
</tr>
<tr>
<td></td>
<td>All ED2 (ATA vs LBI)</td>
<td>46.34% vs 46.57%</td>
<td>$t(12) = 0.279$, p &gt; 0.05</td>
<td>±2.179</td>
<td>$H_0$ cannot be rejected.</td>
</tr>
<tr>
<td>¶ 1 ED1 (ATA vs LBI)</td>
<td>40.06% vs 29.47%</td>
<td>$t(11) = 1.102$, p &gt; 0.05</td>
<td>±2.201</td>
<td>$H_0$ cannot be rejected.</td>
<td></td>
</tr>
<tr>
<td>¶ 1 ED2 (ATA vs LBI)</td>
<td>41.84% vs 59.31%</td>
<td>$t(11) = -.540$, p &gt; 0.05</td>
<td>±2.201</td>
<td>$H_0$ cannot be rejected.</td>
<td></td>
</tr>
<tr>
<td>¶ 2 ED1 (ATA vs LBI)</td>
<td>56.25% vs 37.13%</td>
<td>$t(11) = 0.650$, p &gt; 0.05</td>
<td>±2.201</td>
<td>$H_0$ cannot be rejected.</td>
<td></td>
</tr>
<tr>
<td>¶ 2 ED2 (ATA vs LBI)</td>
<td>51.58% vs 35.65%</td>
<td>$t(11) = 0.945$, p &gt; 0.05</td>
<td>±2.201</td>
<td>$H_0$ cannot be rejected.</td>
<td></td>
</tr>
</tbody>
</table>

Notes: ¶ = paragraph; $H_0$ = the null hypothesis (no significant difference)
The table above demonstrates that based on most of the means, the ATA Framework appears to be more effective than the LBI Bandscale. Five (5) of the means demonstrate that the ATA Framework is more effective than the LBI Bandscale, one means shows a very small difference of effectiveness between the two models, and another one shows that the LBI Bandscale is more effective than the ATA Framework. However, from all the t-test results, there is no significant difference between the two assessment models, which also means the data do not show that there is a large difference between the two models. As a result, the null hypothesis cannot be rejected. In other words, the first hypothesis cannot be supported.

The findings from these statistical calculations are similar to the findings of several Translation Studies scholars described in Chapter 3. Waddington (2001a, 2001b, & 2003), Conde (2009a & 2011), and Turner et al. (2010) discover that two translation assessment models, the error deduction/analysis and the assessment scale/holistic approach, are equally valid as assessment tools. Waddington discovers that an error analysis approach is more reliable than a holistic approach, and in this study most of the means show that the ATA Framework is more effective than the LBI Bandscale. Conde and Turner et al. reveal in their research that both assessment models are equally effective, and in this study based on the t-test the ATA Framework and the LBI Bandscale are not very much different in terms of their effectiveness.

7.1.2 Respondents’ Preferences on Translation Assessment Models

Based on the second hypothesis, the study attempts to reveal students’ preferences on translation assessment models, and in this case it is the choice between the ATA Framework and the LBI Bandscale. To support or reject the second hypothesis, the analysis of the survey
results were designed to provide the answer. Based on the pilot project survey results, all of the respondents who filled in the online survey think that both of these two models are effective. However, the responses from this survey might be biased since the respondents only saw and experienced one type of assessment model.

Based on the formal project survey results, 10 out of 11 respondents filling in the survey prefer the ATA Framework to the LBI Bandscale as the assessment model that can improve their translation skills. One prefers the LBI Bandscale to the ATA Framework. Nevertheless, none of the respondents considers the LBI Bandscale ineffective. To conclude, the second hypothesis can be supported based on the formal or second project survey results, although all respondents still consider both assessment models effective.

7.1.3 Correlation between Respondent Perspectives and Their Translation Results

Based on the third hypothesis, the study attempts to discover the correlation between student perspectives on the assessment models and their translation results. Based on Table 6.4 showing Pilot Project Data and Table 6.7 showing Second Project Data, almost all respondents’ revisions show improvement or a decrease in errors. There are only four outliers, showing no improvement or an increase in errors, out of 26 respondents who completed all the tasks in the study. These outliers are discussed in Chapter 6. In addition, based on the survey results of pilot and second projects, all respondents show positive attitude towards the two assessment models. They all think that the ATA Framework and the LBI Bandscale are effective in helping them improve their translations. In conclusion, there is a positive correlation between student perspectives on these two translation assessment models and their translation results. Thus, the third hypothesis can be supported.


7.1.4 Translation Error Types and Two Assessment Models

In Chapter 6, there is a discussion of translation errors. This discussion produces the definition of translation errors based on the projects of this study, which can be classified as mistranslation, misunderstanding, and unclear, awkward, and ambiguous meanings and/or expressions. Based on the error types in these projects, the most common errors are grammatical errors which are more language errors than translation errors. The second and third common errors, incorrect usage and omission, can be both language and translation errors. The pure translation errors, such as mistranslation/misunderstanding and literalness/faithfulness, do not frequently appear in the translations of the respondents, and they appear even less in the revisions. This phenomenon happened probably because the respondents understood the (Indonesian) source text well, but their English writing still needed to be improved.

In order to discover error types, the use of an error analysis model, such as the ATA Framework, is very fruitful. An assessment scale like the LBI Bandscale cannot be applied to reveal error types as it does not provide the system to do so. In this case, the ATA Framework is more advantageous than the LBI Bandscale to provide details of errors students make. Nonetheless, the use of the LBI Bandscale is not all bad as it provides fixed feedback showing positive and negative points of a translation result. According to Huot, holistic scoring or an assessment scale, such as the LBI Bandscale, is the fastest, most economical model of assessment (1990:239). Thus, the translation teachers in LBI might prefer to use the LBI Bandscale to assess student translations to save time. This can be understood as teachers in Indonesia have to teach many classes, more than five classes a week, in order to maintain adequate income.
To apply just one type of assessment in LBI classes may lead to a dilemma. Most of the students prefer the ATA Framework based on the results of this study, while the teachers may incline to choose the LBI Bandscale as it will not burden them in doing assessment. The fair solution is to apply both assessment models with some modification. For instance, the marking of the errors will apply the ATA Framework along with its weighted scale, and the feedback given to students will use the scale of the LBI Bandscale. However, on the one hand, the feedback in the LBI Bandscale needs to be improved first and adjusted with the error types described in the ATA Framework. On the other hand, several error types in the ATA Framework may require some changes and adjustment with the common error types occurring in the Indonesian-English and English-Indonesian translations. Certainly, further research is necessary to discover the constant translation error types in this language pair beforehand.

7.2 Limitations

There are a number of limitations in this study. First of all, the researcher was not the teacher in the class where the projects were conducted as this was a remote study, so the researcher did not know for sure what was going in the class and what the respondents learned from the class. It was also impossible to communicate directly in person to provide explanations on the two assessment models applied in the study. The respondents were sent seven files containing the explanation of the two models besides the assessment and the feedback for their work. It might not have been easy for the respondents to understand all the explanations of these two assessment models in a short time by just reading about them. Practice on the models is actually required for students to comprehend the way the models work. In classes where the models are discussed in class and used over and over again in
multiple classes, this will produce an environment where the students truly understand the feedback in a more profound way.

Another disadvantage for the researcher not being the teacher is that it is difficult to understand the reason(s) for the outliers appearing in the study. It was impossible to reach out to these outliers to have direct discussion about what happened to them or to try to resolve their issues. Another disadvantage is that there was no two-way communication between the researcher and the respondents despite the fact they were encouraged to ask questions and to contact the researcher anytime if they had problems related to the research. None of the respondents asked any questions or initiated any discussions. They might have been reluctant to do that with an unknown researcher who was far away and whom they had never met. In addition, due to the distance between the researcher and the respondents, the projects could not be conducted in a learner-oriented environment. The study actually aims to support a student-centered classroom by applying the originally summative assessment models, the ATA Framework and the LBI Bandscale, as formative assessment or to assess formative exercises. Unfortunately, the conditions of the remote research provided obstacles in applying either approach.

Another limitation is the small number of respondents from the LBI translation classes even after over two years, which happened because not many people in Indonesia can afford this translator’s training or have interest in learning translation, so the number of the participants who register for these classes is always low. Even if the study had been conducted for over five years, the number of respondents would still have not reached 100 or even 60. Consequently, this present study cannot produce scientific results with only 26 respondents in total who completed all the tasks (only 22 completing the survey, however). The bigger the sample size, the more chance there is to obtain more significant results as increased numbers will more reliably reflect the population mean.
Therefore, the results from this study are merely descriptive explaining the indications of a phenomenon in this particular group of respondents, and they might not be able to be applied in a different sampling of respondents.

Another limitation is that the study could have included Rater 2 already during the translation phase to provide more ‘objective results’ from the beginning. However, involving Rater 2 from the translation phase would have confused and burdened the respondents even more with many tasks to do in this study. Furthermore, since the independent judge or Rater 2 lives and works in Indonesia and he is also busy with his own studies as a PhD student while having classes to teach and being in charge as a coordinator for the LBI translation classes, it would have been too much to impose on him to be involved from the translation phase. This issue was solved with good inter-rater reliability between Rater 1 and Rater 2 as discussed in Chapter 6. Good inter-rater reliability shows that both raters are valid and reliable to assess translations and revisions. It is true that the inter-rater reliability in this study only demonstrates having similar number of errors between the two raters, but to discover whether the two raters also found similar error types is impossible in this case, as Rater 2 applied only the LBI Bandscale. He is not familiar with the ATA Framework or any error analysis system, so he only marked errors without any explanation. If Rater 1 had helped to identify Rater 2’s error types, it means the error types would have been decided all by Rater 1 and this would be biased.

Another limitation is that the researcher neglected to put in the translation brief or in the task instructions for whom the respondents are asked to translate the text. However, considering the level of the respondents’ translation and English skills, it would not make any difference whether they had been informed for whom they were supposed to do the
translations or not, as most of them were struggling with language and transfer errors. In a higher level class, more specific translation brief will probably show a difference in results.

7.3 Future Research

For future research, a longitudinal study conducted to obtain more results of students’ improvement using these two assessment tools would further demonstrate the effectiveness of both translation assessment tools and increase understanding of the use of these tools. Once I return to Indonesia, I will be able to conduct future projects in my own translation classes so that I can personally explore the effectiveness of the assessment tools in class from the beginning to the end of the class term. I will be able to integrate information about the assessment criteria into my classroom discussions so that students have a better understanding of the assessment criteria. I will also be able to introduce student-centered feedback into the classroom discussions to further enhance their abilities to improve their work. I have described in Chapter 5 that the students in the LBI translation classes rarely received written feedback for the formative exercises because the teachers will just discuss the results in class. Based on this study, I can introduce to other LBI fellow teachers the idea of providing written feedback for student translations and revisions. I will also teach the students to provide feedback for each other or peer-feedback using one or both assessment tools applied in this study. Thus, besides continuing the discussion of the exercises in class, students will have their chance to revise their work based on the suggestions given by the teacher and their classmates.

In addition, I will be able to involve other translation teachers by asking them to become raters in the research conducted in my class(es) from the beginning. Those teachers will benefit from becoming good raters using two established assessment translation models,
and they might be encouraged to also conduct research in their classes. Another further research direction is that the same type of study can be conducted from English to Indonesian, thus allowing the students to translate into their native language. The comparison of research results from two different translation directions, Indonesian to English and English to Indonesian, will reveal interesting and maybe important findings, such as similar and different assessment strategies of Indonesian to English translations and English to Indonesian translations.

This pilot observational study has generated indications or speculations on the effectiveness of the ATA Framework and the LBI Bandscale from the participants of the LBI translation classes. These results are not yet scientific results, but with further research in the future the study might be able to provide more comprehensive data and evidence to reveal the effectiveness of these two assessment models, or as noted, to develop another, perhaps more effective approach. This study can provide an example of how feedback should be conducted in a translator training course. As described in previous chapters, the respondents were happy to receive written feedback to improve their translations, and almost all of them did improve. Most importantly, the chance for the students to redo their first translation draft is what makes them really learn to be better, which enables them to become good translators in the future. Discussion of translation exercises in class only is not sufficient without the follow-up by the students to revise their work. Thus, this study also provides a model for the teachers in LBI translation classes specifically and the translation teachers in Indonesia in general to improve their procedures by providing feedback for translation students and by affording students the opportunity to revise their work.
REFERENCES


Kasandrinou, M. 2010. Evaluation as a means of quality assurance: what is it and how can it be done? In V. Pellatt, K. Griffiths, & S-C. Wu (Eds.), *Teaching and testing interpreting and translating*, (pp. 193-205). Frankfurt am Main: Peter Lang Publishing.


Kramer, G. L., Hanson, C., & Olsen, D. 2010. Assessment Frameworks that can make a difference in achieving institutional outcomes. In G.L. Kramer & R.L. Swing (Eds.),
Higher Education Assessments: Leadership Matters (pp. 27-56). Lanham, MD:


*Instructional Science, 18*, 145-165.


Schaefer, K. M., & Zygmont, D. 2003. Analyzing the Teaching Style of Nursing Faculty: Does It Promote a Student-Centered or Teacher-Centered Learning Environment? 
*Nursing Education Perspectives, 24*(5), 238-246.


_Teaching Assistant Orientation_ (powerpoint slides). Retrieved from:


_British Journal of Educational Studies, 53_(4), 466-478.


APPENDIX A

THE LBI BANDSCALE
### ASSESSMENT BANDSCALE FOR PP LBI UI TRANSLATOR TRAININGS

<table>
<thead>
<tr>
<th>Indonesian</th>
<th>English</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>A</strong> Pemahaman TSu dan penulisan TSa baik sekali. Sesekali secara kreatif mampu menemukan padanan yang sangat sesuai.</td>
<td>The source text (ST) understanding and the target text (TT) writing are excellent. Sometimes the student can creatively discover very suitable equivalents.</td>
<td>A = 4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>A- = 3.5</td>
</tr>
<tr>
<td><strong>B</strong> Pemahaman TSu baik, namun adakalanya terjadi kesalahpahaman TSu, terutama jika menerjemahkan bagian teks yang sulit. Penulisan dalam BSa umumnya baik, tidak banyak membuat kesalahan ejaan dan/atau tanda baca.</td>
<td>The ST understanding is good, but occasionally there is ST misunderstanding, especially when translating a difficult part of the text. The writing in the target language (TL) is generally good, not making many errors in spelling and/or punctuation marks.</td>
<td>B+ = 3.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>B = 3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>B- = 2.8</td>
</tr>
<tr>
<td><strong>C</strong> Pemahaman TSu cukup baik jika tingkat kesulitan teks tidak tinggi. Namun jika teks memiliki banyak ungkapan idiomatis atau terminologi khusus, peserta sering tidak mampu memahami teks dengan baik. Dalam hal penulisan dalam BSa, peserta seringkali membuat kesalahan yang terkait dengan pilihan kata, kolokasi, ejaan dan tanda baca.</td>
<td>The ST understanding is quite good if the text difficulty level is not high. However, if the text has a large number of idiomatic expressions or special terminology, the student will often be unable to understand the text well. In terms of the TL writing, the student often makes mistakes related to choices of words, collocation, spelling, and punctuation marks.</td>
<td>C+ = 2.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>C = 2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>C- = 1.5</td>
</tr>
<tr>
<td><strong>D</strong> Pemahaman TSu perlu ditingkatkan lagi. Banyak kesalahan pengungkapan pesan ke dalam BSa yang menyebabkan salah pengertian.</td>
<td>The ST understanding needs to be improved further. Many errors in the message transfer into the TL causing misunderstanding.</td>
<td>D = 1</td>
</tr>
</tbody>
</table>
APPENDIX B

THE ATA FRAMEWORK
# ATA Framework for Error Marking

<table>
<thead>
<tr>
<th>One</th>
<th>Two</th>
<th>Four</th>
<th>Eight</th>
<th>Sixteen</th>
<th>Code</th>
<th>Reason</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>MT</td>
<td>- Mistranslation into target language</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>MU</td>
<td>- Misunderstanding of the source text</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>A</td>
<td>- Addition</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>O</td>
<td>- Omission</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>T</td>
<td>- Terminology, word choice</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>R</td>
<td>- Register</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>F</td>
<td>- Faithfulness</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>L</td>
<td>- Literalness</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>FA</td>
<td>- Faux ami (false friend)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>COH</td>
<td>- Cohesion</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>AMB</td>
<td>- Ambiguity</td>
</tr>
</tbody>
</table>
|     |     |      |       |         | ST   | - Style
|     |     |      |       |         | IND  | - Indecision, gave more than one option |
|     |     |      |       |         | INC  | - Inconsistency |
|     |     |      |       |         | CLS  | - Failure to use items mentioned in class |
|     |     |      |       |         | WO   | - Word Order |

### Translation/strategic/transfer errors: Negative impact on understanding/use of target text

### Mechanical errors: Negative impact on overall quality of target text.

| G   | Grammar
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>SYN</td>
<td>Syntax (phrase/clause/sentence structure)</td>
</tr>
<tr>
<td>P</td>
<td>Punctuation</td>
</tr>
<tr>
<td>SP</td>
<td>Spelling</td>
</tr>
<tr>
<td>D</td>
<td>Diacritics</td>
</tr>
<tr>
<td>C</td>
<td>Capitalization</td>
</tr>
<tr>
<td>WF</td>
<td>Word form (part of speech)</td>
</tr>
<tr>
<td>U</td>
<td>Usage</td>
</tr>
</tbody>
</table>

Total Error Points:

Quality Points: 0

Explanation: 0

Combined Score: 0

---

1Style can include collocation, zeugma, and denominalization errors.

2Grammar can include agreement
### ATA Framework for Standard Error Marking — Explanations

<table>
<thead>
<tr>
<th>Code</th>
<th>Criteria and Description of Each Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>INC</td>
<td><strong>Incomplete Passage:</strong> Missing titles, headings, or sentences within a passage may be marked as one or more errors of omission, depending on how much is omitted.</td>
</tr>
<tr>
<td>ILL</td>
<td><strong>Illegible:</strong> ATA criterion not used in KSU context because our texts are never handwritten.</td>
</tr>
</tbody>
</table>
| MU   | **Misunderstanding of Original Text:** This category applies when the grader can see—usually by back-translating the target-language text—that the error arises from misreading a word, for example, or misinterpreting the syntax of a sentence. In other words, the result is wrong because the translation was based on a misunderstood source text. Example: *Die große Versuchung, die die chinesischen Rechtümer auf die Hunnen ausübten…*  
Correct: The great temptation that the Chinese riches exerted on the Huns…  
Wrong: The great experiment [Versuch] that the Chinese kingdoms [Reiche] practiced upon the Huns…  
Example: The deal's complexity is mind-bending, involving six countries charged with putting down no fewer than nine guerrilla outfits that operate inside Congo. Here, the phrase charged with means responsible for, not accused of. |
<p>| MT   | <strong>Mistranslation into Target Language:</strong> The meaning of the original text is not conveyed properly in the target language. For example, a term in the translated text might be much more general (<em>scientists</em> instead of <em>researchers</em>; <em>protein</em> instead of <em>albumin</em>) or more specific (<em>stallion</em> instead of <em>horse</em>) than the original term. Mistranslations can also involve the choice of prepositions, the use of definite and indefinite articles, and the choice of verb tense and mood. |
| A/O  | <strong>Addition or Omission:</strong> Something is inserted that is not clearly expressed in the original text, or something essential to the meaning is left out. Explicitation and implicitation are acceptable, however, but should be well justified (e.g., in your log and during class discussion.) |
| T MJ | <strong>Terminology, Word Choice:</strong> This error often involves terms used in various technical, legal, and financial contexts, where words often have very specific meanings. In more general texts, the candidate might not have selected the most appropriate word among several that have similar (but not identical) meanings. Sometimes a very specific word in general language is precisely right for a given situation due to usage, collocational conventions, etc. (mot juste). |
| R    | <strong>Register:</strong> Language level, degree of formality should be preserved in the translation; examples of errors include using everyday words instead of medical terms (<em>spit</em> instead of <em>saliva</em>), making a legal document sound journalistic, <em>du/Sie</em> errors, anachronisms, and culturally inappropriate expressions. |
| F    | <strong>Too Freely Translated:</strong> Candidates are asked to translate the meaning and intent of the source text, not to rewrite or improve upon it. The grader will carefully compare the translation to the source text. If a “creative” rendition changes the meaning, an error will be marked. If recasting a sentence—i.e., altering the order of its major elements—destroys the flow, changes the emphasis, or obscures the author’s intent, an error may be marked. |
| L    | <strong>Too Literal, Word-for-Word:</strong> Translations that follow the source text exactly may result in awkward, often incorrect renditions. Translate literally when it works, but not at the expense of clarity and natural syntax. |</p>
<table>
<thead>
<tr>
<th>Code #</th>
<th>Criteria and Description of Each Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>FC</td>
<td><strong>False Cognate:</strong> False cognates, sometimes called false friends or <em>faux amis</em>, are words that probably are derived from similar roots and that have very similar or identical forms, but that have different meanings, at least in some contexts. Example: aktuell ≠ actually, but rather current.</td>
</tr>
<tr>
<td>IND</td>
<td><strong>Indecision, Giving More than One Option:</strong> You must choose between possible options, even if we have stated in class that more than one option is acceptable. You cannot leave the reader to choose between options.</td>
</tr>
<tr>
<td>I</td>
<td><strong>Inconsistency, Same Term Translated Differently:</strong> In general, a term that is used consistently in the source text should be translated consistently into the target language. Conversely, if the source text uses different words for the same idea interchangeably, the candidate should try to come up with a similar variety in the target language. [Note: this is not always the case in scientific and technical texts, where a proliferation of synonyms in the source language should be treated with consistent terminology in English.]</td>
</tr>
<tr>
<td>AMB</td>
<td><strong>Ambiguity:</strong> If the meaning is clear in the source text but ambiguous in the translation, an error may be marked. The reader should not have to puzzle out the meaning.</td>
</tr>
<tr>
<td>G</td>
<td><strong>Grammar:</strong> Grammatical errors include lack of agreement between subject and verb, incorrect verb forms, incorrect case of nouns, pronouns, or adjectives, and use of an adjective where an adverb is needed.</td>
</tr>
</tbody>
</table>
| SYZ   | **Syntax (Phrase/Clause/Sentence Structure):** The arrangement of words or other elements of a sentence should conform to the rules of the target language. Errors in this category include sentence fragments, improper modification, lack of parallelism, and unnatural word order. If incorrect syntax changes or obscures the meaning, the error is more serious. Examples:  
  *The cupboard was infested with moths that organic foods enjoy.* (Nonsense)  
  *Climbing the mountain, the cabin came into view.* (Dangling participle)  
  *She skis, plays tennis, and rock climbing.* (Zugma, parallelism error)  
  *The from Berlin coming and now on platform 12 arriving express train is two hours late.* (Interference) |
<p>| P     | <strong>Punctuation:</strong> The conventions of the target language should be followed, including those governing the use of quotation marks, commas, semicolons, and colons. Incorrect or unclear paragraphing is counted as an error. |
| SP    | <strong>Spelling:</strong> A spelling error that causes confusion about the intended meaning is more serious (e.g., principle/principal, systemic/systematic, peddle/pedal, dear/deer, bear/bare, sight/site. Context is a factor as well. |
| D     | <strong>Accents and Other Diacritical Marks:</strong> The conventions of the target language should be followed consistently. If incorrect or missing diacritical marks obscure the meaning, the error is more serious. This is, of course, rarely a problem for into—English translation, but one does need to know when to retain accents (e.g., résumé) and when not. Aberrant accents are an irritation (<em>engâge</em> for <em>engagé</em>, the term used for hired help or mercenaries during the voyages of exploration in the early Louisiana Territory; <em>Là Place</em>, a shopping center in Beechwood). |
| C     | <strong>Case (Upper/Lower):</strong> The conventions of the target language should be followed. Examples: Don't carry German capitalization of nouns over into English, or the non-capitalization of the first words in lists, etc. |</p>
<table>
<thead>
<tr>
<th>Code #</th>
<th>Criteria and Description of Each Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>WF</td>
<td><strong>Word Form:</strong> The root of the word is correct, but the wrong form is used. Example in English: The product has been tampered with and is no longer safety.</td>
</tr>
<tr>
<td>U</td>
<td><strong>Usage:</strong> Correct and idiomatic usage of the target language is expected. Errors include the use of the wrong preposition or misuse of a grammatical form. Examples: take vs. make a walk, married to vs. married with, etc.</td>
</tr>
<tr>
<td>ST</td>
<td><strong>Style:</strong> If the source text is characterized by a distinctive manner of expression—flowery, staccato, conversational, instructional—this should be reflected in the translation. Awkward or clumsy renditions that obscure the meaning may also be penalized.</td>
</tr>
<tr>
<td>COL</td>
<td><strong>Collocation:</strong> Two or more frequently used words that can be consecutive or non-consecutive, that form a unit of meaning, and that are accepted by common usage. Ex: The representative <em>made</em> a speech. (It is not proper to ‘hold’ a speech, or to ‘perform’ a speech.)</td>
</tr>
<tr>
<td>WO</td>
<td><strong>Word Order:</strong> The relative order of subject, object, and verb, of modifiers (adverbs adjectives, numerals, demonstratives, possessives, etc.) and of clauses and other syntactic elements used in any given language. Note: Word order tends to differ from language to language and requires modification during translation.</td>
</tr>
<tr>
<td>D</td>
<td><strong>Denominalization:</strong> A translation procedure where a noun or nominal structure from the source text is transformed into a verbal structure in the target text. Note: Some languages prefer nouns, and others prefer verbs.</td>
</tr>
</tbody>
</table>

*Based in part on Doyle, *The ATA Chronicle* | November/December 2003; ATA Website, [http://www.atanet.org/bin/view.pl/12438.html](http://www atanet.org/bin/view.pl/12438.html)*
FLOWCHART FOR ERROR POINT DECISIONS
Version 2009

Understanding / usefulness of target text affected? Meaning lost / obscured?

no = mechanical error (identifiable without viewing source text); maximum 4 points

yes = translation / transfer / strategy error

See overall questions, below at left.

Effect on understanding / use / content negligible?

yes

0 points

no

Error apparent to any editor?

no

0 points

yes

1 point

Error apparent to typical target reader? (text still intelligible)

no

1 point

yes

2 points

Error disruptive to typical target reader? (effort required for understanding)

no

2 points

yes

4 points

Is interference minimal?

yes

2 points

no

Is disruption limited in scope?

yes

4 points

no

4 points

Text as a whole still usable despite this serious error?

yes

8 points

no

16 points

Overall questions to guide decisions:
Can target text be used for intended purpose?
Is it intelligible to the intended target reader?
Does it transfer the meaning of the source text?
APPENDIX C

FLYERS AND CONSENT FORMS
The Flyer of the Pilot Projects
(and Its Translation into Indonesian)

TRANSLATION ASSESSMENT WITH DETAILED FEEDBACK

Principal Investigator: Dr. Sue Ellen Wright
Co-Investigator: Haru Deliana Dewi

Background:
Empirical research on translation assessment is rare worldwide, and almost non-existent in Indonesia. Thus, it is difficult to find evidence on the effectiveness of assessment given to translation students. Therefore, this research will attempt to compare the effectiveness of different styles of feedback provided for translation assessment and to reveal student perspectives on translation assessment.

Procedure:
Before we begin with the research, you will be asked to sign a consent form if you are willing to participate in the study. The study will ask you to work on a translation at home. The text to be translated will be a general text consisting of around 150 to 200 words. You will be expected to work individually and in one hour. After you complete the first assignment, you should directly send it to hdewi@kent.edu.

Two weeks later, I will send back your work with assessment and feedback. You will be requested to read the assessment and feedback carefully as you will be asked to do a revision of your first version. You will have a week to do the revision. After you are finished with the revision and send it to hdewi@kent.edu, you will be asked to fill in a simple short survey.

I understand that your language proficiency in both English and Indonesia is more than sufficient. However, as one of the requirements from my university, I must provide the translation of all research instruments to prevent misunderstanding from occurring. Thus, you will find that all materials are written both in English and Indonesia.

Risks and benefits:
There will be no risks for you as the participant since the participation will be voluntary and anonymous. You can discontinue your participation at any time without any penalty. Your identity will be kept confidential and will not be mentioned in the report. Your performance in this research will have no impact on your grades in class.

The potential benefits of participating in this study include having your translation work assessed and providing you with feedback to help improve your translation skills. Once you complete your participation, you will be awarded with a token of appreciation, such as a keychain or something similar, and a certificate of participation sent directly from Kent State University after the research is concluded. In addition, you will receive a short report of the research result.
PENILAIAN TERJEMAHAN DENGAN UMPAN BALIK TERPERINCI

Peneliti Utama: Dr. Sue Ellen Wright  
Peneliti Kedua: Haru Deliana Dewi

Latar Belakang:

Prosedur:

Dua minggu kemudian, saya akan mengirimkan kembali hasil kerja anda dengan penilaian dan umpan balik. Anda akan diminta membaca penilaian dan umpan balik tersebut secara seksama karena anda akan diminta melakukan revisi atas versi terjemahan pertama anda. Anda akan mempunyai waktu selama seminggu untuk melakukan revisi. Setelah anda selesai dengan revisi terjemahan anda dan mengirimkannya ke hdewi@kent.edu, anda akan diminta mengisi survei pendek yang sederhana.

Saya memahami bahwa kemahiran bahasa anda baik dalam bahasa Inggris dan dalam bahasa Indonesia lebih dari cukup. Namun, sebagai salah satu syarat dari universitas saya, saya harus menyediakan terjemahan dari semua instrumen/materi riset untuk mencegah kesalahpahaman terjadi. Jadi, anda akan menemukan bahwa semua materi ditulis dalam bahasa Inggris dan bahasa Indonesia.

Risiko dan keuntungan:
Tidak akan ada risiko bagi anda sebagai peserta riset karena partisipasi bersifat sukarela dan tidak mengungkapkan identitas anda. Anda dapat menghentikan partisipasi anda kapan saja tanpa adanya penalti. Identitas anda akan dirahasiakan dan tidak akan disebut dalam laporan hasil riset. Kinerja anda dalam riset ini tidak akan mempunyai dampak pada nilai anda di kelas.

Keuntungan potensial berpartisipasi dalam penelitian ini termasuk dinilainya kerja terjemahan anda dan memberikan anda umpan balik untuk membantu memperbaiki keahlian menerjemahkan anda. Setelah anda menyelesaikan partisipasi anda, anda akan diberikan kenang-kenangan, seperti gantungan kunci atau sejenisnya, dan sertifikat partisipasi yang langsung dikirim dari Kent State University sesudah penelitian ini selesai. Selain itu, anda akan menerima laporan pendek berisi hasil riset.
The Flyer of the Formal or Second Projects
(and Its Translation into Indonesian)

COMPARING TWO TRANSLATION ASSESSMENT MODELS: CORRELATING
STUDENT REVISIONS AND PERSPECTIVES

Principal Investigator: Dr. Sue Ellen Wright
Co-Investigator: Haru Deliana Dewi, M.Hum.

Background:
Worldwide there are a number of ways that translator trainers evaluate student translation, but very little empirical research exists to judge their effectiveness. Thus, it is difficult to find evidence on the effectiveness of assessment given to translation students. Therefore, this research will attempt to compare the effectiveness of two different translation assessment models. We also want to give you the opportunity to tell us how you feel about the styles of assessment we will use.

Procedure:
Before we begin with the research, you will be asked to sign a consent form if you are willing to participate in the study. The study will ask you to work on a translation at home. The text to be translated will be a general text consisting of around 314 words. You will be expected to work individually and in up to three (3) hours (of course, you may very well finish sooner than this). After you complete the first assignment, you should directly send it to hdewi@kent.edu.

Two weeks later, you will receive your work back with two different styles of assessment. You will be requested to read the assessment and feedback carefully as you will be asked to do a revision of your first version. You will have a week to do the revision. After you are finished with the revision and send it to hdewi@kent.edu, you will be asked to fill in a simple short survey.

I understand that your language proficiency in both English and Indonesian is more than sufficient. However, as one of the requirements from my university, I must provide the translation of all research instruments to prevent misunderstanding from occurring. Thus, you will find that all materials are written both in English and Indonesia.

Risks and benefits:
There will be no risks for you as the participant since the participation will be voluntary and anonymous. You can discontinue your participation at any time without any penalty. Your identity will be kept confidential and will not be mentioned in the report. Your performance in this research will have no impact on your grades in class.

The potential benefits of participating in this study include having your translation work assessed and providing you with feedback to help improve your translation skills. Once you complete your participation, you will be awarded with a token of appreciation, such as a keychain or something similar, and a certificate of participation sent directly from Kent State University after the research is concluded. In addition, you will receive a short report of the research result.
MEMBANDINGKAN DUA MODEL PENILAIAN TERJEMAHAN:
MENGKORELASIKAN REVISI TERJEMAHAN DENGAN PANDANGAN
PESERTA PENERJEMAHAN

Penyelidik Utama: Dr. Sue Ellen Wright
Penyelidik Kedua: Haru Deliana Dewi, M.Hum.

Latar Belakang:

Prosedur:
Sebelum memulai riset ini, anda akan diminta menandatangani suatu surat perkenalan tertulis apabila anda bersedia berpartisipasi dalam penelitian ini. Penelitian ini akan meminta anda mengerjakan suatu terjemahan di rumah anda sendiri. Teks yang diterjemahkan berupa teks umum terdiri atas sekitar 314 kata. Anda diharapkan bekerja sendiri dan dalam waktu tiga (3) jam (tentu saja anda dapat menyelesaikannya kurang dari 3 jam). Setelah anda menyelesaikan tugas pertama, anda diminta mengirimkannya ke hdewi@kent.edu. Dua minggu kemudian, anda akan menerima hasil kerja anda dengan dua model penilaian yang berbeda. Anda akan diminta membaca penilaian dan umpan balik tersebut secara seksama karena anda akan diminta melakukan revisi atas versi terjemahan pertama anda. Anda akan mempunyai waktu selama seminggu untuk melakukan revisi. Setelah anda selesai dengan revisi terjemahan anda dan mengirimkannya ke hdewi@kent.edu, anda akan diminta mengisi survei pendek yang sederhana.

Saya memahami bahwa kemahiran bahasa anda baik dalam bahasa Inggris dan dalam bahasa Indonesia lebih dari cukup. Namun, sebagai salah satu syarat dari universitas saya, saya harus menyediakan terjemahan dari semua instrumen/materi riset untuk mencegah kesalahpahaman terjadi. Jadi, anda akan menemukan bahwa semua materi ditulis dalam bahasa Inggris dan bahasa Indonesia.

Risiko dan keuntungan:
Tidak akan ada risiko bagi anda sebagai peserta riset karena partisipasi bersifat sukarela dan tidak mengungkapkan identitas anda. Anda dapat menghentikan partisipasi anda kapan saja tanpa adanya penalti. Identitas anda akan dirahasiakan dan tidak akan disebut dalam laporan hasil riset. Kinerja anda dalam riset ini tidak akan mempunyai dampak pada nilai anda di kelas.

Keuntungan potensial berpartisipasi dalam penelitian ini termasuk dinilainya kerja terjemahan anda dan memberikan anda umpan balik untuk membantu memperbaiki keahlian menerjemahkan anda. Setelah anda menyelesaikan partisipasi anda, anda akan diberikan kenang-kenangan, seperti gantungan kunci atau sejenisnya, dan sertifikat partisipasi yang
langsung dikirim dari Kent State University sesudah penelitian ini selesai. Selain itu, anda akan menerima laporan pendek berisi hasil riset.
The Consent Form of the Pilot Projects
(and Its Translation into Indonesian)

Informed Consent to Participate in a Research Study

Study Title: Translation Assessment with Detailed Feedback
Principal Investigator: Dr. Sue Ellen Wright
Co-Investigator: Haru Deliana Dewi

You are being invited to participate in a research study. This consent form will provide you with information on the research project, what you will need to do, and the associated risks and benefits of the research. Your participation is entirely voluntary. Please, read this form carefully. It is important that you ask questions and fully understand the research in order to make an informed decision. You will receive a copy of this document to take with you.

Purpose: This study aims to discover the effectiveness of translation assessment with detailed feedback and to reveal student perspectives on translation assessment.

Procedure: You will be presented with a short source text and be requested to translate it into Indonesian or English (depending on the course in which you are enrolled) at home. You will be expected to work individually for just one hour and can use any resources available. After you have finished, please send the results to hdewi@kent.edu. Two weeks later, you will receive the assessment and feedback on your work, and you will be requested to revise your first version based on this assessment and feedback. You will have a week to do the revision. Please also send the revision to hdewi@kent.edu. Then you will be asked to fill in a simple short survey to report your perspective on that assessment.

Benefits: The potential benefits of participating in this study include having your translation assessed and receiving feedback to help improve your translation skills. Once you have completed your participation, you will be entitled to receive a token of appreciation, such as a keychain or something similar, and a certificate of participation directly from Kent State University after the research is concluded. You will also be provided with a short report describing the results of this research, which may provide insights into your own understanding of translation assessment in the future.

Risks and Discomforts: There are no anticipated risks beyond those encountered in everyday life.

Privacy and Confidentiality: No identifying information will be collected. Your signed consent form will be kept separate from your study data, and responses/results will not be linked to you.

Voluntary Participation: Taking part in this research study is entirely voluntary. You may choose not to participate or you may discontinue your participation at any time without penalty or loss of benefits to which you are otherwise entitled. You will be informed of any
new, relevant information that may affect your health, welfare, or willingness to continue your study participation.

**Contact Information:** If you have any questions or concerns about this research, you may contact Haru Deliana Dewi at hdewi@kent.edu. This project has been approved by the Kent State University Institutional Review Board. If you have any questions about your rights as a research participant or complaints about the research, you may call the IRB at +1 330-672 1797 or kmccreal@kent.edu.

(The accompanying Indonesian form is a true and fair equivalent of this original English form. Kent State University requires that all materials presented to subjects be translated into their native language. See the third and fourth pages of the Indonesian version of this document.)

I have read this consent form and have had the opportunity to have my questions answered to my satisfaction. I voluntarily agree to participate in this study. I understand that a copy of this consent will be provided to me for future reference.

(Saya telah membaca formulir persetujuan tertulis ini dan telah mempunyai kesempatan bagi pertanyaan saya terjawab secara memuaskan. Saya secara sukarela setuju untuk berpartisipasi dalam penelitian ini. Saya mengerti bahwa salinan persetujuan tertulis ini akan disediakan bagi saya untuk rujukan di masa depan.)

________________________________________
Participant Signature Date
*(Tanda Tangan Peserta) (Tanggal)*
Persetujuan Tertulis untuk Berpartisipasi dalam Suatu Penelitian

Judul Penelitian: Penilaian Terjemahan dengan Umpan Balik Terperinci
Penyelidik Utama: Dr. Sue Ellen Wright
Penyelidik Kedua: Haru Deliana Dewi


Tujuan: Penelitian ini bertujuan untuk mengetahui keefektifan penilaian terjemahan dengan umpan balik terperinci dan mengungkapkan pandangan peserta pelatihan penerjemahan mengenai penilaian penerjemahan.

Prosedur: Anda akan diberikan teks sumber pendek dan diminta untuk menerjemahkannya ke dalam bahasa Indonesia atau bahasa Inggris (tergantung pada pelatihan yang anda ikuti) di rumah masing-masing. Anda akan diharapkan untuk bekerja hanya selama satu jam secara individu dan dapat memanfaatkan segala sumber bantuan tersedia, seperti kamus, internet, dll. Setelah anda selesai mengerjakan, mohon kirim hasilnya ke email (hdewi@kent.edu). Dua minggu kemudian, anda akan menerima penilaian dan umpan balik atas hasil terjemahan anda, dan kemudian anda akan diminta untuk merevisi tugas pertama anda berdasarkan penilaian dan umpan balik yang anda dapat. Anda akan mempunyai waktu selama satu minggu untuk melakukan revisi. Mohon kirimkan juga revisi tersebut ke email peneliti. Selanjutnya, anda akan diminta untuk mengisi survei sederhana yang pendek untuk mengetahui pandangan anda akan penilaian yang diberikan.

Keuntungan: Potensi keuntungan berpartisipasi dalam penelitian ini adalah hasil terjemahan anda akan dinilai dan diberikan umpan balik untuk membantu memperbaiki keahlian penerjemahan anda. Setelah selesai berpartisipasi dalam riset ini, anda berhak diberikan suatu kenang-kenangan, seperti gantungan kunci atau sejenisnya, dan sertifikat partisipasi langsung dari Kent State University sesudah penelitian ini selesai. Anda juga akan dikerahkan laporan pendek dari hasil penelitian, yang akan memberikan pandangan terhadap pemahaman anda sendiri akan penilaian penerjemahan di masa depan.

Risiko dan Ketidaknyamanan: Tidak akan ada risiko yang diantisipasi melebihi risiko yang biasa dihadapi dalam kehidupan sehari-hari.

Privasi dan Kerahasiaan: Tidak akan ada informasi dengan identitas yang akan dikumpulkan. Formulir persetujuan tertulis anda akan disimpan secara terpisah dari data penelitian anda, dan segala jawaban/hasil tidak akan dihubungkan ke anda.

Partisipasi secara Sukarela: Berpartisipasi dalam riset ini sepenuhnya bersifat sukarela. Anda dapat memilih untuk tidak berpartisipasi atau anda dapat menghentikan partisipasi anda kapan saja tanpa adanya penalti atau hilangnya keuntungan yang seharusnya berhak anda miliki. Anda akan diberitahukan segala informasi baru yang relevan yang mungkin
mempengaruhi kesehatan, kesejahteraan, atau kerelaan anda untuk melanjutkan partisipasi anda dalam riset ini.

Informasi Kontak: Apabila anda mempunyai pertanyaan atau kekhawatiran mengenai riset ini, anda dapat menghubungi Haru Deliana Dewi pada hdewi@kent.edu. Proyek ini telah disetujui oleh IRB (Lembaga Peninjau Institusi) Kent State University. Apabila anda mempunyai pertanyaan mengenai hak anda sebagai peserta penelitian atau keluhan mengenai riset ini, anda dapat menghubungi IRB di 330-672 1797 atau kmcrea1@kent.edu
Informed Consent to Participate in a Research Study

Study Title:
COMPARING TWO TRANSLATION ASSESSMENT MODELS: CORRELATING STUDENT REVISIONS AND PERSPECTIVES

Principal Investigator: Dr. Sue Ellen Wright
Co-Investigator: Haru Deliana Dewi, M.Hum.

You are being invited to participate in a research study. This consent form will provide you with information on the research project, what you will need to do, and the associated risks and benefits of the research. Your participation is entirely voluntary. Please, read this form carefully. It is important that you ask questions and fully understand the research in order to make an informed decision. You will receive a copy of this document to take with you.

Purpose: This study aims to discover the effectiveness of two translation assessment models and to correlate student perspectives on assessment with their translation results.

Procedure: You will be presented with a short source text (2 paragraphs, each consisting of around 150 to 160 words) and be requested to translate them into English at home. You will be expected to work individually for three (3) hours and can use any resources available. After you have finished, please send the results to hdewi@kent.edu. Two weeks later, you will receive the assessment and feedback on your work. One paragraph will be assessed using one model of assessment, and the other paragraph will be assessed using the other model of assessment. Next, you will be requested to revise your first versions based on this assessment and feedback. You will have a week to do the revision. Please also send the revisions to hdewi@kent.edu. Then you will be asked to fill in a simple short survey to report your perspective on these two models of assessment.

Benefits: The potential benefits of participating in this study include having your translation assessed and receiving feedback to help improve your translation skills. Once you have completed your participation, you will be entitled to receive a token of appreciation, such as a keychain or something similar, and a certificate of participation directly from Kent State University after the research is concluded. You will also be provided with a short report describing the results of this research, which may provide insights into your own understanding of translation assessment in the future.

Risks and Discomforts: There are no anticipated risks beyond those encountered in everyday life.

Privacy and Confidentiality: No identifying information will be collected. Your signed consent form will be kept separate from your study data, and responses/results will not be linked to you.

Voluntary Participation: Taking part in this research study is entirely voluntary. You may choose not to participate or you may discontinue your participation at any time without penalty or loss of benefits to which you are otherwise entitled. You will be informed of any new, relevant information that may affect your health, welfare, or willingness to continue your study participation.

Contact Information: If you have any questions or concerns about this research, you may contact Haru Deliana Dewi at hdewi@kent.edu. This project has been approved by the Kent State University
Institutional Review Board. If you have any questions about your rights as a research participant or complaints about the research, you may call the IRB at +1 330-672 1797 or kmccrea1@kent.edu.

(The accompanying Indonesian form is a true and fair equivalent of this original English form. Kent State University requires that all materials presented to subjects be translated into their native language. See the third and fourth pages of the Indonesian version of this document.)

I have read this consent form and have had the opportunity to have my questions answered to my satisfaction. I voluntarily agree to participate in this study. I understand that a copy of this consent will be provided to me for future reference.

(Saya telah membaca formulir persetujuan tertulis ini dan telah mempunyai kesempatan bagi pertanyaan saya terjawab secara memuaskan. Saya secara sukarela setuju untuk berpartisipasi dalam penelitian ini. Saya mengerti bahwa salinan persetujuan tertulis ini akan disediakan bagi saya untuk rujukan di masa depan.)

Participant Signature
(Tanda Tangan Peserta)  Date
(Tanggal)
Persetujuan Tertulis untuk Berpartisipasi dalam Suatu Penelitian

Judul Penelitian: Membandingkan Dua Model Penilaian Terjemahan: Mengkorelasikan Revisi Terjemahan dengan Pandangan Peserta Penerjemahan

Penyelidik Utama: Dr. Sue Ellen Wright
Penyelidik Kedua: Haru Deliana Dewi, M.Hum.


Tujuan: Penelitian ini bertujuan untuk mengetahui keefektifan dua model penilaian terjemahan dan mengkorelasikan pandangan peserta mengenai penilaian terjemahan dengan hasil terjemahan mereka.


Keuntungan: Potensi keuntungan berpartisipasi dalam penelitian ini adalah hasil terjemahan anda akan dinilai dan diberikan umpan balik untuk membantu memperbaiki keahlian penerjemahan anda. Setelah selesai berpartisipasi dalam riset ini, anda berhak diberikan suatu kenang-kenangan, seperti gantungan kunci atau sejenisnya, dan sertifikat partisipasi langsung dari Kent State University sesudah penelitian ini selesai. Anda juga akan dikirimkan laporan pendek dari hasil penelitian, yang akan memberikan pandangan terhadap pemahaman anda sendiri akan penilaian penerjemahan di masa depan.

Risiko dan Ketidaknyamanan: Tidak akan ada risiko yang diantisipasi melebihi risiko yang biasa dihadapi dalam kehidupan sehari-hari.

Privasi dan Kerahasiaan: Tidak akan ada informasi dengan identitas yang akan dikumpulkan. Formulir persetujuan tertulis anda akan disimpan secara terpisah dari data penelitian anda, dan segala jawaban/hasil tidak akan dihubungkan ke anda.

Partisipasi secara Sukarela: Berpartisipasi dalam riset ini sepenuhnya bersifat sukarela. Anda dapat memilih untuk tidak berpartisipasi atau anda dapat menghentikan partisipasi anda kapan saja tanpa adanya penalti atau hilangnya keuntungan yang seharusnya berhak anda miliki. Anda akan diberitahukan segala informasi baru yang relevan yang mungkin mempengaruhi kesehatan, kesejahteraan, atau kerelaan anda untuk melanjutkan partisipasi anda dalam riset ini.
Informasi Kontak: Apabila anda mempunyai pertanyaan atau kekhawatiran mengenai riset ini, anda dapat menghubungi Haru Deliana Dewi pada hdewi@kent.edu. Proyek ini telah disetujui oleh IRB (Lembaga Peninjau Institusi) Kent State University. Apabila anda mempunyai pertanyaan mengenai hak anda sebagai peserta penelitian atau keluhan mengenai riset ini, anda dapat menghubungi IRB di 330-672 1797 atau kmccrea1@kent.edu
APPENDIX D

FIRST AND FINAL ASSIGNMENT INSTRUCTIONS
The First Assignment Instructions of the Pilot Projects
(and Its Translation into Indonesian)

INSTRUCTIONS FOR THE FIRST ASSIGNMENT:
TRANSLATION FROM ENGLISH TO INDONESIAN
(INSTRUKSI UNTUK TUGAS PERTAMA:
PENERJEMAHAN DARI BAHASA INGGRIS KE BAHASA INDONESIA)

Instructions for the first assignment (instruksi untuk tugas pertama):
Translation of a short general text from English to Indonesian (Penerjemahan teks umum pendek dari bahasa Inggris ke bahasa Indonesia)

- Your task will be to translate the yellow highlighted part of the following document in the space provided below the source text. (Tugas anda adalah menerjemahkan bagian berwarna kuning dari dokumen di bawah ini dan menerjemahkannya di bawah teks sumber.)
- You may have only one hour to complete the task. (Anda dapat menyelesaikan tugas ini hanya selama satu jam.)
- You may use any resources, such as dictionaries, internet, and others. (Anda diperbolehkan menggunakan segala alat bantu penerjemahan, seperti kamus, internet, dan lain-lain.)
- Please work individually. (Mohon bekerja sendiri-sendiri.)
- Do not consult anyone else for help. (Jangan bekerja sama dengan orang lain dalam menerjemahan.)
- Your file is already named based on your complete name and the type of assignment, so please do not change the name of this file. (Berkas anda ini telah diberi nama berdasarkan nama lengkap anda dan tipe tugas, jadi mohon jangan mengubah nama berkas ini.)
- For example, if your name is Anton Fermana and you are doing a translation, then the file name will be antonfermana_translation.doc. (Misalnya, jika nama anda Anton Fermana dan anda mengerjakan penerjemahan, maka nama berkas ini menjadi antonfermana_translation.doc.)
- After you have finished, please submit the work to hdewi@kent.edu. (Setelah anda selesai, mohon kirim hasilnya ke hdewi@kent.edu.)

English source text: (Teks Sumber Berbahasa Inggris:)

Eurozone unemployment reaches new record high in April

The BBC's Jamie Robertson says the employment figures show "disparity across Europe".

Unemployment in the eurozone has reached another record high, according to official figures.
The seasonally-adjusted rate for April was 12.2%, up from 12.1% the month before.

An extra 95,000 people were out of work in the 17 countries that use the euro, taking the total to 19.38 million.

Both Greece and Spain have jobless rates above 25%. The lowest unemployment rate is in Austria at 4.9%.

The European Commission's statistics office, Eurostat, said Germany had an unemployment rate of 5.4% while Luxembourg's was 5.6%.

The highest jobless rates are in Greece (27.0% in February 2013), Spain (26.8%) and Portugal (17.8%).

In France, Europe's second largest economy, the number of jobless people rose to a new record high in April.

"We do not see a stabilisation in unemployment before the middle of next year," said Frederik Ducrozet, an economist at Credit Agricole in Paris. "The picture in France is still deteriorating."

'Social crisis'

Youth unemployment remains a particular concern. In April, 3.6 million people under the age of 25 were out of work in the eurozone, which translated to an unemployment rate of 24.4%.

Figures from the Italian government showed 40.5% of young people in Italy are unemployed.
**ECB action?**

Some consider that the ECB needs to do more than simply cutting interest rates to boost economic activity and create jobs.

Earlier this month, the ECB lowered its benchmark interest rate to 0.50% from 0.75%, the first cut in 10 months, and said it was "ready to act if needed" if more measures were required to boost the eurozone's economic health.

In its report earlier this week, the OECD hinted that the ECB might want to expand quantitative easing (QE) as a measure to encourage stronger growth.

Nick Matthews, a senior economist at Nomura International, said: "We do not expect a strong recovery in the eurozone.

"It puts pressure on the ECB to deliver even more conventional and non-conventional measures," he added.

The European Central Bank is due to meet next week.

(source: http://www.bbc.co.uk/news/business-22727373)

**The translation:** (Terjemahannya:)
INSTRUCTIONS FOR THE FIRST ASSIGNMENT:
TRANSLATION FROM INDONESIAN TO ENGLISH
(INSTRUKSI UNTUK TUGAS PERTAMA:
PENERJEMAHAN DARI BAHASA INDONESIA KE BAHASA INGGRIS)

Instructions for the first assignment (instruksi untuk tugas pertama):
Translation of a short general text from Indonesian to English (Penerjemahan teks umum pendek dari bahasa Indonesia ke bahasa Inggris)

- Your task will be to translate the yellow highlighted part of the following document in the space provided below the source text. (Tugas anda adalah menerjemahkan bagian berwarna kuning dari dokumen di bawah ini dan menerjemahkannya di bawah teks sumber.)
- You may have only one hour to complete the task. (Anda dapat menyelesaikan tugas ini hanya selama satu jam.)
- You may use any resources, such as dictionaries, internet, and others. (Anda diperbolehkan menggunakan segala alat bantu penerjemahan, seperti kamus, internet, dan lain-lain.)
- Please work individually. (Mohon bekerja sendiri-sendiri.)
- Do not consult anyone else for help. (Jangan bekerja sama dengan orang lain dalam menerjemahan.)
- Your file is already named based on your complete name and the type of assignment, so please do not change the name of this file. (Berkas anda ini telah diberi nama berdasarkan nama lengkap anda dan tipe tugas, jadi mohon jangan mengubah nama berkas ini.)
- For example, if your name is Anton Fermana and you are doing a translation, then the file name will be antonfermana_translation.doc. (Misalnya, jika nama anda Anton Fermana dan anda mengerjakan penerjemahan, maka nama berkas ini menjadi antonfermana_translation.doc.)
- After you have finished, please submit the work to hdewi@kent.edu. (Setelah anda selesai, mohon kirim hasilnya ke hdewi@kent.edu.)

Indonesian Source Text: (Teks Sumber Berbahasa Indonesia)
Tembus Rekor Lagi, Pengangguran di 17 Negara Pengguna Euro Capai 19,38 Juta
Wahyu Daniel - detikfinance
Sabtu, 01/06/2013 10:36 WIB

London - Jumlah pengangguran di 17 negara Eropa pengguna mata uang euro kembali menembus rekor. Pada April 2013, persentase pengangguran di 17 negara tersebut mencapai 12,2%, naik dari bulan sebelumnya 10%.

Dikutip dari BBC, Sabtu (1/6/2013), ada tambahan pengangguran sebanyak 95 ribu orang di April 2013. Sehingga total pengangguran di 17 negara Eropa tersebut mencapai 19,38 juta orang.
Angka pengangguran tertinggi terjadi di Yunani dan Spanyol yang persentasenya 25%. Sedangkan angka pengangguran terendah adalah di Austria sebesar 4,9%.

Kantor Komisi Statistik Eropa yaitu Eurostat mengatakan, angka pengangguran di Jerman adalah 5,4%, sementara di Luxemburg adalah 5,6%. Angka pengangguran di Yunani adalah 27%, di Spanyol 26,8%, dan Portugal 17,8%.


Pengangguran remaja juga terus meningkat. Jumlah pengangguran berusia di bawah 25 tahun pada April 2013 berjumlah 3,6 juta orang, atau persentasenya 24,4%.

Jumlah pengangguran remaja terbesar terjadi di Italia dengan angka 40,5%. Dalam setahun terakhir yang berakhir April 2013, ada 1,6 juta orang kehilangan pekerjaannya di 17 negara Eropa tersebut.

(dnl/dnl)


The Translation: (Terjemahannya:)
The Final Assignment Instructions of the Pilot Projects
(and Its Translation into Indonesian)

INSTRUCTIONS FOR THE FINAL ASSIGNMENT
(INSTRUKSI UNTUK TUGAS AKHIR)

Instructions for the final submission: (instruksi untuk tugas akhir:)
Revision of the first assignment (revisi tugas pertama)

- Please find your assessment and feedback. Study the assessment carefully. (Mohon lihat penilaian dan umpan balik anda. Pelajari penilaianannya secara seksama.)
- Please revise your initial assignment based on the assessment and feedback given to you. (Mohon revisi tugas pertama anda berdasarkan penilaian dan umpan balik yang diberikan kepada anda.)
- Copy your first assignment in the space provided below to be revised there. (Salin tugas pertama anda pada tempat yang telah disediakan di bawah ini untuk direvisi.)
- Please work individually. (Mohon bekerja sendiri-sendiri.)
- Do not consult anyone else for help. (Jangan bekerja sama dengan orang lain dalam merevisi tugas anda.)
- You will have one week to complete the task. (Anda akan mempunyai waktu satu minggu untuk menyelesaikan tugas ini.)
- Your file is already named based on your complete name and the type of assignment, so please do not change the name of this file. (Berkas anda telah diberi nama berdasarkan nama lengkap anda dan tipe tugas, jadi mohon tidak mengubah nama berkas ini.)
- For example, if your name is Anton Fermana and you are doing a revision, then the file name will be antonfermana_revision.doc. (Misalnya, jika nama anda Anton Fermana dan anda mengerjakan revisi, maka nama berkas ini menjadi antonfermana_revision.doc.)
- After you have finished, please send the revision to hdewi@kent.edu. (Setelah anda selesai, mohon kirim revisi anda ke hdewi@kent.edu.)
- When you have sent the revision, please go http://www.surveymonkey.com/r/GRGJZG8 to take a short survey. (Ketika anda telah mengirimkan revisi anda, mohon mengisi survei pendek dengan mengklik situs berikut ini http://www.surveymonkey.com/r/GRGJZG8.)

The revision: (Revisinya:)
The First Assignment Instructions of the Formal or Second Projects
(and Its Translation into Indonesian)

INSTRUCTIONS FOR THE FIRST ASSIGNMENT:
TRANSLATION FROM INDONESIAN TO ENGLISH
(INSTRUKSI UNTUK TUGAS PERTAMA:
PENERJEMAHAN DARI BAHASA INDONESIA KE BAHASA INGGRIS)

Instructions for the first assignment (instruksi untuk tugas pertama):
Translation of a short general text from Indonesian to English (Penerjemahan teks umum pendek dari bahasa Indonesia ke bahasa Inggris)

- Your task will be to translate the following two paragraphs in the space provided below the source text. (Tugas anda adalah menerjemahkan dua paragraf di bawah ini dan menerjemahkannya di bawah teks sumber.)
- You may have up to three (3) hour to complete the task, but you don’t have to take that long. (Anda dapat menyelesaikan tugas ini selama tiga (3) jam, tetapi anda tidak diharuskan menyelesaikan selama itu.)
- You may use any resources, such as dictionaries, internet, and others. (Anda diperbolehkan menggunakan segala alat bantu penerjemahan, seperti kamus, internet, dan lain-lain.)
- Please work individually. (Mohon bekerja sendiri-sendiri.)
- Do not consult anyone else for help. (Jangan bekerja sama dengan orang lain dalam menerjemahan.)
- Your file is already named based on your complete name and the type of assignment, so please do not change the name of this file. Later your name will be removed from the data so that your participation will become anonymous. (Berkas anda ini telah diberi nama berdasarkan nama lengkap anda dan tipe tugas, jadi mohon jangan mengubah nama berkas ini. Nanti nama anda akan dihapus dari data sehingga partisipasi anda akan tanpa nama.)
- For example, if your name is Anton Fermana and you are doing a translation, then the file name will be antonfermana_translation.doc. (Misalnya, jika nama anda Anton Fermana dan anda mengerjakan penerjemahan, maka nama berkas ini menjadi antonfermana_translation.doc.)
- After you have finished translating both Paragraph 1 and Paragraph 2, please submit the work to hdewi@kent.edu. (Setelah anda selesai, mohon kirim hasilnya ke hdewi@kent.edu.)

186
Paragraph 1:

*Berita tentang mahalnya biaya sekolah.* Pembaca yang menjadi peserta FGD berpendapat bahwa biaya RSBI sebenarnya tidak semahal sekolah swasta. Namun demikian, biaya yang tidak seragam antar sekolah dan ditetapkan oleh masing-masing sekolah ini yang kemudian membuka peluang bagi sekolah untuk menentukan harga semaunya. Sebagian peserta FGD mengatakan bahwa isi berita tentang mahalnya biaya dinyatakan memberikan referensi yang baik sehingga mereka mencari tahu dan membandingkan antar sekolah RSBI. Ketidak seragaman biaya menurut mereka tidak baik. Seharusnya pemerintah membuat keputusan tentang penyeragaman biaya sesuai dengan taraf hidup masyarakat di mana RSBI tersebut berada. Terhadap isi berita mengenai mahalnya biaya memunculkan kesenjangan dan menutup peluang bagi keluarga kurang mampu untuk menyekolahkan anaknya di RSBI. Ketidak seragaman biaya menurut mereka tidak baik. Seharusnya pemerintah membuat keputusan tentang penyeragaman biaya sesuai dengan taraf hidup masyarakat di mana RSBI tersebut berada. Terhadap isi berita mengenai mahalnya biaya memunculkan kesenjangan dan menutup peluang bagi keluarga kurang mampu untuk menyekolahkan anaknya di RSBI. Terhadap isi berita mengenai mahalnya biaya memunculkan kesenjangan dan menutup peluang bagi keluarga kurang mampu untuk menyekolahkan anaknya di RSBI. Terhadap isi berita mengenai mahalnya biaya memunculkan kesenjangan dan menutup peluang bagi keluarga kurang mampu untuk menyekolahkan anaknya di RSBI. Terhadap isi berita mengenai mahalnya biaya memunculkan kesenjangan dan menutup peluang bagi keluarga kurang mampu untuk menyekolahkan anaknya di RSBI. Terhadap isi berita mengenai mahalnya biaya memunculkan kesenjangan dan menutup peluang bagi keluarga kurang mampu untuk menyekolahkan anaknya di RSBI. Terhadap isi berita mengenai mahalnya biaya memunculkan kesenjangan dan menutup peluang bagi keluarga kurang mampu untuk menyekolahkan anaknya di RSBI. Terhadap isi berita mengenai mahalnya biaya memunculkan kesenjangan dan menutup peluang bagi keluarga kurang mampu untuk menyekolahkan anaknya di RSBI. Terhadap isi berita mengenai mahalnya biaya memunculkan kesenjangan dan menutup peluang bagi keluarga kurang mampu untuk menyekolahkan anaknya di RSBI. Terhadap isi berita mengenai mahalnya biaya memunculkan kesenjangan dan menutup peluang bagi keluarga kurang mampu untuk menyekolahkan anaknya di RSBI. Terhadap isi berita mengenai mahalnya biaya memunculkan kesenjangan dan menutup peluang bagi keluarga kurang mampu untuk menyekolahkan anaknya di RSBI. Terhadap isi berita mengenai mahalnya biaya memunculkan kesenjangan dan menutup peluang bagi keluarga kurang mampu untuk menyekolahkan anaknya di RSBI. Terhadap isi berita mengenai mahalnya biaya memunculkan kesenjangan dan menutup peluang bagi keluarga kurang mampu untuk menyekolahkan anaknya di RSBI. Terhadap isi berita mengenai mahalnya biaya memunculkan kesenjangan dan menutup peluang bagi keluarga kurang mampu untuk menyekolahkan anaknya di RSBI. Terhadap isi berita mengenai mahalnya biaya memunculkan kesenjangan dan menutup peluang bagi keluarga kurang mampu untuk menyekolahkan anaknya di RSBI. Terhadap isi berita mengenai mahalnya biaya memunculkan kesenjangan dan menutup peluang bagi keluarga kurang mampu untuk menyekolahkan anaknya di RSBI. Terhadap isi berita mengenai mahalnya biaya memunculkan kesenjangan dan menutup peluang bagi keluarga kurang mampu untuk menyekolahkan anaknya di RSBI. Terhadap isi berita mengenai mahalnya biaya memunculkan kesenjangan dan menutup peluang bagi keluarga kurang mampu untuk menyekolahkan anaknya di RSBI. Terhadap isi berita mengenai mahalnya biaya memunculkan kesenjangan dan menutup peluang bagi keluarga kurang mampu untuk menyekolahkan anaknya di RSBI. Terhadap isi berita mengenai mahalnya biaya memunculkan kesenjangan dan menutup peluang bagi keluarga kurang mampu untuk menyekolahkan anaknya di RSBI. Terhadap isi berita mengenai mahalnya biaya memunculkan kesenjangan dan menutup peluang bagi keluarga kurang mampu untuk menyekolahkan anaknya di RSBI. Terhadap isi berita mengenai mahalnya biaya memunculkan kesenjangan dan menutup peluang bagi keluarga kurang mampu untuk menyekolahkan anaknya di RSBI. Terhadap isi berita mengenai mahalnya biaya memunculkan kesenjangan dan menutup peluang bagi keluarga kurang mampu untuk menyekolahkan anaknya di RSBI. Terhadap isi berita mengenai mahalnya biaya memunculkan kesenjangan dan menutup peluang bagi keluarga kurang mampu untuk menyekolahkan anaknya di RSBI. Terhadap isi berita mengenai mahalnya biaya memunculkan kesenjangan dan menutup peluang bagi keluarga kurang mampu untuk menyekolahkan anaknya di RSBI. Terhadap isi berita mengenai mahalnya biaya memunculkan kesenjangan dan menutup peluang bagi keluarga kurang mampu untuk menyekolahkan anaknya di RSBI. Terhadap isi berita mengenai mahalnya biaya memunculkan kesenjangan dan menutup peluang bagi keluarga kurang mampu untuk menyekolahkan anaknya di RSBI. Terhadap isi berita mengenai mahalnya biaya memunculkan kesenjangan dan menutup peluang bagi keluarga kurang mampu untuk menyekolahkan anaknya di RSBI. Terhadap isi berita mengenai mahalnya biaya memunculkan kesenjangan dan menutup peluang bagi keluarga kurang mampu untuk menyekolahkan anaknya di RSBI. Terhadap isi berita mengenai mahalnya biaya memunculkan kesenjangan dan menutup peluang bagi keluarga kurang mampu untuk menyekolahkan anaknya di RSBI. Terhadap isi berita mengenai mahalnya biaya memunculkan kesenjangan dan menutup peluang bagi keluarga kurang mampu untuk menyekolahkan anaknya di RSBI. Terhadap isi berita mengenai mahalnya biaya memunculkan kesenjangan dan menutup peluang bagi keluarga kurang mampu untuk menyekolahkan anaknya di RSBI. Terhadap isi berita mengenai mahalnya biaya memunculkan kesenjangan dan menutup peluang bagi keluarga kurang mampu untuk menyekolahkan anaknya di RSBI. Terhadap isi berita mengenai mahalnya biaya memunculkan kesenjangan dan menutup peluang bagi keluarga kurang mampu untuk menyekolahkan anaknya di RSBI. Terhadap isi berita mengenai mahalnya biaya memunculkan kesenjangan dan menutup peluang bagi keluarga kurang mampu untuk menyekolahkan anaknya di RSBI. Terhadap isi berita mengenai mahalnya biaya memunculkan kesenjangan dan menutup peluang bagi keluarga kurang mampu untuk menyekolahkan anaknya di RSBI. Terhadap isi berita mengenai mahalnya biaya memunculkan kesenjangan dan menutup peluang bagi keluarga kurang mampu untuk menyekolahkan anaknya di RSBI. Terhadap isi berita mengenai mahalnya biaya memunculkan kesenjangan dan menutup peluang bagi keluarga kurang mampu untuk menyekolahkan anaknya di RSBI. Terhadap isi berita mengenai mahalnya biaya memunculkan kesenjangan dan menutup peluang bagi keluarga kurang mampu untuk menyekolahkan anaknya di RSBI. Terhadap isi berita mengenai mahalnya biaya memunculkan kesenjangan dan menutup peluang bagi keluarga kurang mampu untuk menyekolahkan anaknya di RSBI. Terhadap isi berita mengenai mahalnya biaya memunculkan kesenjangan dan menutup peluang bagi keluarga kurang mampu untuk menyekolahkan anaknya di RSBI. Terhadap isi berita mengenai mahalnya biaya memunculkan kesenjangan dan menutup peluang bagi keluarga kurang mampu untuk menyekolahkan anaknya di RSBI. Terhadap isi berita mengenai mahalnya biaya memunculkan kesenjangan dan menutup peluang bagi keluarga kurang mampu untuk menyekolahkan anaknya di RSBI. Terhadap isi berita mengenai mahalnya biaya memunculkan kesenjangan dan menutup peluang bagi keluarga kurang mampu untuk menyekolahkan anaknya di RSBI. Terhadap isi berita mengenai mahalnya biaya memunculkan kesenjangan dan menutup peluang bagi keluarga kurang mampu untuk menyekolahkan anaknya di RSBI. Terhadap isi berita mengenai mahalnya biaya memunculkan kesenjangan dan menutup peluang bagi keluarga kurang mampu untuk menyekolahkan anaknya di RSBI. Terhadap isi berita mengenai mahalnya biaya memunculkan kesenjangan dan menutup peluang bagi keluarga kurang mampu untuk menyekolahkan anaknya di RSBI. Terhadap isi berita mengenai mahalnya biaya memunculkan kesenjangan dan menutup peluang bagi keluarga kurang mampu untuk menyekolahkan anaknya di RSBI. Terhadap isi berita mengenai mahalnya biaya memunculkan kesenjangan dan menutup peluang bagi keluarga kurang mampu untuk menyekolahk
Paragraph 2:
Berita tentang penggunaan bahasa Inggris sebagai bahasa pengantar. Pembaca berita menyikapi hal ini dengan cukup bijak. Artinya, sebenarnya kebijakan pemerintah dinilai benar bahwa hanya sebagian dari mata pelajaran yang menggunakan bahasa pengantar bahasa Inggris. Namun demikian, sekolah yang justru dianggap memaksakan diri menggunakan bahasa pengantar bahasa Inggris untuk semua mata pelajaran. Selain gurunya tidak semua mampu dan siap, demikian pula siswanya yang belum merata kemampuan bahasa Inggrisnya. Meski demikian, sebagian besar peserta mengatakan setuju bahwa sebaiknya semua mata pelajaran diberikan dalam bahasa Indonesia dan penguatan bahasa Inggris diberikan pada mata pelajaran bahasa Inggris itu sendiri dengan ditambah berbagai aktivitas penunjang. Penggunaan bahasa Inggris dinilai akan kontraproduktif dengan penanaman nilai-nilai keindonesiaan yang seharusnya dapat berakar kuat dalam diri siswa, dan menjadi nilai keutamaan yang unggul dalam persaingan global sumber daya manusia nantinya.

The Translation: (Terjemahannya:)}
The Final Assignment Instructions of the Formal or Second Projects
(and Its Translation into Indonesian)

Instructions for the final submission: (instruksi untuk tugas akhir)
Revision of the first assignment (revisi tugas pertama)

- Please find your assessment and feedback. Study the assessment carefully. (Mohon lihat penilaian dan umpan balik anda. Pelajari penilaiannya secara seksama.)
- Please revise your initial assignment based on the assessment and feedback given to you. (Mohon revisi tugas pertama anda berdasarkan penilaian dan umpan balik yang diberikan kepada anda.)
- You will note that there are two models for translation assessment used in this research: the ATA Framework, and the LBI Bandscale. One paragraph has been assessed using one assessment model and the other paragraph has been assessed using the other assessment model. (Anda perlu perhatikan bahwa ada dua model penilaian terjemahan yang digunakan dalam riset ini: ATA Framework dan LBI Bandscale. Satu paragraf dinilai dengan satu model penilaian dan paragraf lainnya dinilai dengan model penilaian lainnya.)
- Copy your first assignment in the space provided below and revise it there. (Salin tugas pertama anda pada tempat yang telah disediakan di bawah ini untuk direvisi.)
- Please work individually. (Mohon bekerja sendiri-sendiri.)
- Do not consult anyone else for help. (Jangan bekerja sama dengan orang lain dalam merevisi tugas anda.)
- You will have one week to complete the task. (Anda akan mempunyai waktu satu minggu untuk menyelesaikan tugas ini.)
- Your file is already named based on your complete name and the type of assignment, so please do not change the name of this file. (Berkas anda telah diberi nama berdasarkan nama lengkap anda dan tipe tugas, jadi mohon tidak mengubah nama berkas ini.)
- For example, if your name is Anton Fermana and you are doing a revision, then the file name will be antonfermana_revision.doc. (Misalnya, jika nama anda Anton Fermana dan anda mengerjakan revisi, maka nama berkas ini menjadi antonfermana_revision.doc.)
- Please note that the names are only used at this initial point in the project so that we can continue to email with one another. After that, your part of the project will be “anonymized”, so that no one will have any way of connecting you with the project. (Mohon diperhatikan bahwa nama-nama hanya digunakan pada tahap awal riset ini agar kami dapat melanjutkan mengirimkan email pada anda. Setelah tahap ini, semua hasil tugas anda akan dihapus namanya sehingga tidak ada seorang pun yang dapat menghubungkan anda dengan riset ini.)
- After you have finished, please send the revision to hdewi@kent.edu. (Setelah anda selesai, mohon kirim revisi anda ke hdewi@kent.edu.)
• After you have sent the revision, please go to https://www.surveymonkey.com/r/J6LHRTL to take a short survey. (Ketika anda telah mengirimkan revisi anda, mohon mengisi survei pendek dengan mengklik situs berikut ini https://www.surveymonkey.com/r/J6LHRTL.)

• This survey is also anonymous. We will not be able to connect your survey comment with your specific translation work. (Survei ini bersifat tanpa nama. Kami tidak akan dapat menghubungkan komentar anda dalam survei dengan hasil terjemahan anda.)

The revision: (Revisinya:)}
APPENDIX E
THE SURVEYS
1. Which type of assessment did you receive?
   A. Group 1
   B. Group 2

2. The assessment and feedback I was given were….
   A. very general
   B. general
   C. specific
   D. very specific

3. The assessment and feedback I was given were….
   A. very simple
   B. simple
   C. complicated
   D. very complicated

4. The assessment and feedback I was given were….
   a. Yes (if yes, answer number 5 and skip number 6)
   b. No (if no, skip number 5 and answer number 6)

5. I can use the assessment and feedback to improve my translation in the following ways:

6. The assessment and feedback are not helpful for me because:
Second Project Survey
Comparing Two Translation Assessment Models

1. What is the first paragraph assessed with?
   A. The ATA Framework
   B. The LBI Bandscale

2. What is the second paragraph assessed with?
   A. The ATA Framework
   B. The LBI Bandscale

3. How difficult is it for you to understand and apply the assessment using the ATA Framework? (choose just one answer!)
   1 = very simple
   2 = simple
   3 = normal
   4 = complicated
   5 = very complicated

4. How difficult is it for you to understand and apply the assessment using the LBI Bandscale? (choose just one answer!)
   1 = very simple
   2 = simple
   3 = normal
   4 = complicated
   5 = very complicated
5. How effective is the assessment using the ATA Framework in helping you revise your translation? (choose just one answer!)
   1 = a great deal less effective
   2 = somewhat less effective
   3 = effective
   4 = somewhat more effective
   5 = a great deal more effective

6. How effective is the assessment using the LBI Bandscale in helping you revise your translation? (choose just one answer!)
   1 = a great deal less effective
   2 = somewhat less effective
   3 = effective
   4 = somewhat more effective
   5 = a great deal more effective

7. Which method of assessment do you prefer for your translation assessment, the ATA Framework or the LBI Bandscale? Why?
APPENDIX F

ILR (The Interagency Language Roundtable) SCALES
Preface These Skill Level Descriptions are primarily intended to serve as guidelines for use in government settings. They are separate and distinct from the ILR Language Skill Level Descriptions for Speaking, Listening, Reading, and Writing.

Translation is the process of transferring text from one language into another. It is a complex skill requiring several abilities. Consequently, extreme care must be exercised in hiring translators or assigning translation tasks to them. To do otherwise entails the risk that imprecise or even wrong information will be conveyed.

The term “translation” is normally reserved for written renditions of written materials. Translation is thereby distinct from interpretation, which produces a spoken equivalent between two languages. While there are correspondences between translation and interpretation skills, the following applies only to document-to-document renderings.

A successful translation is one that conveys the explicit and implicit meaning of the source language into the target language as fully and accurately as possible. From the standpoint of the user, the translation must also meet the prescribed specifications and deadlines.

Competence in two languages is necessary but not sufficient for any translation task. Though the translator must be able to (1) read and comprehend the source language and (2) write comprehensibly in the target language, the translator must also be able to (3) choose the equivalent expression in the target language that both fully conveys and best matches the meaning intended in the source language (referred to as congruity judgment).

A weakness in any of these three abilities will influence performance adversely and have a negative impact on the utility of the product. Therefore, all three abilities must be considered when assessing translation skills.

The prerequisite language skills for translation (reading in the source language and writing in the target language) are tested separately from translation skills themselves. Language proficiency testing serves as a screening tool and may provide a preliminary indication of the individual’s potential as a translator. However, to assess translation skills, a translation test that measures the individual’s ability to exercise congruity judgment and apply a translation methodology successfully must be used.

Various non-linguistic factors have an impact on performance, such as the time allotted to deliver the product. Familiarity with the subject matter and the socio-cultural aspects of either or both source and target languages may also affect performance. Given previous knowledge of these factors or appropriate training, an
individual with limited skills may be able in certain instances to produce renditions of various texts that might be useful for specific purposes. On the other hand, an otherwise skilled translator who lacks subject matter knowledge or who is unfamiliar with certain socio-cultural aspects may provide an unreliable translation of some points if he or she has no access to relevant resources.

Moreover, analytical and research skills as well as adeptness in using translation tools and resources (such as monolingual dictionaries and glossaries, on-line aids, and consultation with experts) allow the individual to proceed methodically and verify the appropriateness of the equivalents chosen. Such specialized skills must be acquired through training and practice.

The complexity of the translation task increases with the complexity of the text to be translated. Accordingly, the individual’s performance range will depend on the degree to which competence in all the necessary skills combine in order to produce a rendition that is useful to a reader not familiar with the source language. Hence, the need for product review and oversight diminishes as the performance level rises. Nonetheless, any translation of potential importance should, to the extent possible, be subject to review for accuracy by another qualified individual.

In summary, an individual’s translation performance level depends on (1) command of two languages, (2) ability to exercise congruity judgment and apply a translation methodology, (3) familiarity with the cultural context of both languages, (4) knowledge of terminology in specialized fields, and (5) ability to finalize the product within time constraints and according to specifications.

Individuals should be assigned to tasks within their performance range. To facilitate this correspondence, the Skill Level Descriptions for Translation are divided into three bands: Minimal Performance (levels 0+ to 1+), Limited Performance (levels 2 and 2+), and Professional Performance (levels 3 to 5).

Examples of tasks and texts appropriate for each level are provided. Each level implies control of all functions at the lower levels. The “plus level” designation is used to describe performance which substantially exceeds the next lower skill level but for any reason does not fully meet the criteria for the next higher level.

The Minimal Performance Levels (0+ to 1+) are characterized by weaknesses in all of the requisite skills. Translation is not possible at these levels.

The Limited Performance Levels (2 to 2+) are characterized by weaknesses in one or more of the requisite skills. Renditions prepared by individuals at these levels should not be considered professional translations, and should be subject to rigorous quality control with feedback as a means of development toward professional status.

It is at the Professional Performance Level 3 that all necessary skills begin to align and enable production of a reasonably accurate and reliable translation.
At Professional Performance Level 4 (and above) an individual’s competence and expertise combine to produce an accurate and reliable translation of a variety of texts ranging from simple to complex.

It must be noted that language tasks often associated with translation, such as gisting and summarizing a text, are not included in this document, since such tasks require skills distinct from translation skills.

Level 5 (Professional Performance):

Can successfully translate virtually all texts, including those where lack of linguistic and cultural parallelism between the source language and the target language requires precise congruity judgments and the ability to apply a translation methodology. Expression is flawless. At this level, the translator consistently excels in a number of specialties, and is generally regarded as one of the arbiters of translating very high level language by persons competent in dealing with such material. Nonetheless, the resulting product may be subject to quality control.

Level 4+ (Professional Performance Plus):

Can successfully apply a translation methodology to translate texts that contain highly original and special purpose language (such as that contained in religious sermons, literary prose, and poetry). At this level, a successful performance requires not only conveying content and register but also capturing to the greatest extent all nuances intended in the source document. Expression is virtually flawless. Can produce fully accurate translations in a number of subject fields. When the need arises to perform in areas outside of specialization, a translator at this level is able to reach a successful level of performance given the time necessary for acquiring the relevant knowledge of the subject matter. The resulting product is a professional translation which may be subject to quality control.

Level 4 (Professional Performance):

Can successfully apply a translation methodology to translate a wide variety of complex texts that contain difficult, abstract, idiomatic, highly technical, and colloquial writing. Able to capture subtleties, nuances, and tone and register (such as official, formal, and informal writing). Such texts range from commentary reflecting a specific culture to analysis and argumentation. Linguistic knowledge and familiarity with source language norms enable an individual at this level to translate handwritten documents and other texts that represent spontaneous expression characteristic of native speakers. Expression reflects native usage and consistent control of target language conventions. Can translate materials outside the individual’s specialties, but may not reach the absolute subject matter accuracy of the specialist in the given field. The resulting product is a professional translation which may be subject to quality control.

Level 3+ (Professional Performance):
Can generally translate a variety of texts, such as many scientific or financial reports, some legal documents and some colloquial writings. Can convey the meaning of many socio-cultural elements embedded in a text as well as most nuances and relatively infrequent lexical and syntactic items of the source language. Expression reflects target language norms and usage. May be able to operate in fields outside areas of specialty. The resulting product is a draft translation, subject to quality control.

**Level 3 (Professional Performance):**

Can translate texts that contain not only facts but also abstract language, showing an emerging ability to capture their intended implications and many nuances. Such texts usually contain situations and events which are subject to value judgments of a personal or institutional kind, as in some newspaper editorials, propaganda tracts, and evaluations of projects. Linguistic knowledge of both the terminology and the means of expression specific to a subject field is strong enough to allow the translator to operate successfully within that field. Word choice and expression generally adhere to target language norms and rarely obscure meaning. The resulting product is a draft translation, subject to quality control.

**Level 2+ (Limited Performance):**

Can render straightforward texts dealing with everyday matters that include statements of fact as well as some judgments, opinion, or other elements which entail more than direct exposition, but do not contain figurative language, complicated concepts, complex sentence structures, or instances of syntactic or semantic skewing. In these types of texts, the individual can read source language materials and render them accurately into the target language, conveying the key points and/or main ideas, supporting facts, most of the details, and some nuances. Can usually operate in more than one narrowly defined subject field, using both linguistic knowledge of the languages involved and familiarity with the subject matter. A tendency to adhere to source language structures may result in target language expressions that may appear to be correct but are awkward or perhaps unidiomatic. Such expressions may sometimes obscure meaning. The resulting product is not a professional translation and must be subject to quality control.

**Level 2 (Limited Performance):**

Able to render into the target language some straightforward, factual texts in the standard variety of the source language. Can typically render accurately uncomplicated prose (such as that used in short identification documents, simple letters, instructions, and some narrative reports) that does not contain figurative language, complex sentence structures, embedding, or instances of syntactic or semantic skewing. Can normally rely on knowledge of the subject matter to operate within one given subject field, consisting of a narrow body of material that is routine, repetitive, and often predictable. Expression in the target language may be faulty, frequently reflecting the structure and word order of the source language. To the extent that faulty expression may obscure or distort meaning, accuracy will suffer. The resulting product is not a professional translation and must be submitted to quality control.
Level 1+ (Minimal Performance):

Able to scan source language texts for specific categories, topics, key points and/or main ideas, generally rendering an accurate report on these but often missing supporting facts and details. Can to some extent render factual materials, such as records or database entries, often relying on real-world knowledge or familiarity with the subject matter. Oversight and review of the product are necessary.

Level 1 (Minimal Performance):

Able to make word by word transfers, not always with accuracy. May be able to identify documents by their label or headings and scan graphic materials, such as charts and diagrams, for items of specific interest. Constant oversight and review of the product are necessary.

Level 0+ (Minimal Performance)

Able to transfer very little information from one language into another, usually representing isolated words and/or phrases. Accuracy is haphazard. Constant oversight is required.

Level 0 (Performance)

Has no practical ability to transfer information from one language into another.
APPENDIX G

IELTS (The International English Language Testing System) SCALES
IELTS | Institutions - IELTS band scores

There is no pass or fail in IELTS. Candidates are graded on their performance, using scores from 1 to 9 for each part of the test – Listening, Reading, Writing and Speaking. The results from the four parts then produce an Overall Band Score.

This unique 9-band system measures scores in a consistent manner – wherever and whenever the test is taken. It is internationally recognised and understood, giving you a reliable international currency. IELTS scores have a recommended validity period of two years.

The IELTS 9-band scale

Each band corresponds to a level of English competence. All parts of the test and the Overall Band Score can be reported in whole and half bands, eg 6.5, 7.0, 7.5, 8.0.

**Band 9: Expert user:** has fully operational command of the language: appropriate, accurate and fluent with complete understanding.

**Band 8: Very good user:** has fully operational command of the language with only occasional unsystematic inaccuracies and inappropriacies. Misunderstandings may occur in unfamiliar situations. Handles complex detailed argumentation well.

**Band 7: Good user:** has operational command of the language, though with occasional inaccuracies, inappropriacies and misunderstandings in some situations. Generally handles complex language well and understands detailed reasoning.

**Band 6: Competent user:** has generally effective command of the language despite some inaccuracies, inappropriacies and misunderstandings. Can use and understand fairly complex language, particularly in familiar situations.

**Band 5: Modest user:** has partial command of the language, coping with overall meaning in most situations, though is likely to make many mistakes. Should be able to handle basic communication in own field.

**Band 4: Limited user:** basic competence is limited to familiar situations. Has frequent problems in understanding and expression. Is not able to use complex language.

**Band 3: Extremely limited user:** conveys and understands only general meaning in very familiar situations. Frequent breakdowns in communication occur.

**Band 2: Intermittent user:** no real communication is possible except for the most basic information using isolated words or short formulae in familiar situations and to meet immediate needs. Has great difficulty understanding spoken and written English.

**Band 1: Non-user:** essentially has no ability to use the language beyond possibly a few isolated words.

**Band 0: Did not attempt the test:** No assessable information provided.
APPENDIX H

NAATI (The National Accreditation Authority for Translators and Interpreters)
Master of Translating & Interpreting and Graduate Diploma in Translating and Interpreting

Supplementary Program Information on NAATI accreditation

CONTENTS

1. NAATI Accreditation Requirements

2. Format And Assessment Of The Accreditation Examination

3. Examination Protocols and Procedures

4. How To Graduate and Apply For NAATI Accreditation
Section 1: NAATI ACCREDITATION REQUIREMENTS

1.1 The Master of Translating & Interpreting and Graduate Diploma in Translating and Interpreting are approved by the National Accreditation Authority for Translators and Interpreters (NAATI) at the Professional Level (formerly 'Level 3').

1.2 NAATI accreditation requirements therefore apply to some assessment tasks in Interpreting Accrediting Practice and Translating Accrediting Practice in the Program (see relevant Course Guide about which assessment tasks are also assessed for NAATI accreditation).

1.3 To be eligible for the RMIT academic award, students must PASS all courses of the Program. An RMIT Pass means a minimum mark/grade of 50%. Individual Course Guides have details about the number and type of Assessment Tasks, and the distribution and weighting of marks for each Assessment Task.

1.4 In addition to receiving the award of Master of Translating and Interpreting or Graduate Diploma in Translating and Interpreting, a student who performs at NAATI accreditation standards in certain assessment tasks (refer Course Guide to see which assessment tasks) in Translation Accrediting Practice and/or Interpreting Accrediting Practice, the student will be recommended to NAATI for accreditation.

N.B. A student must complete the masters or the graduate diploma before it can recommend the student to NAATI for accreditation(s).

Section 2: FORMAT AND ASSESSMENT OF THE ACCREDITATION EXAMINATION

2.1 Translation Accrediting Practice (one language direction)

The assessment task used for NAATI recommendation is referred to as Accreditation Examination and consists of:

Written translation of two passages out of three, of 250 words, in the prescribed language direction in two different subject areas, to be done in one sitting of 2.5 hours. The passages may not be the same as the ones practised in class and may include complex language which an educated speaker of the language would be expected to be familiar with. Topics of the three passages are chosen randomly and students will not be informed of them beforehand.
Note students who study translation in the direction from Chinese into English will be given 4 passages, two in traditional script and two in simplified script. Students must choose one passage from the two in traditional scripts and one from the two in simplified scripts.

Each passage is marked out of 50. Please refer Appendix 1 for NAATI marking sheet and RMIT marking descriptors.

To be recommended to NAATI for accreditation, students must be graded by two independent examiners and achieve an average mark of 70% for the 2 passages, with a minimum mark of 32.5 out of 50 for any single passage.

2.2 Interpreting Accrediting Practice (both language directions):

The assessment task used for NAATI recommendation is referred to as Accreditation Examination and consists of three components:

- **Two bilingual dialogues** in two different subject areas. Each dialogue is approximately 400 words long, with roughly 200 words in the English and Language Other Than English (LOTE) parts.

- **Two sight translation passages**, approximately 200 words each relating to the topics of the dialogues. Sight translation done after Dialogue 1 relates to the topic of Dialogue 1 and is from English into LOTE. Sight Translation 2 done after Dialogue 2 relates to the topic of Dialogue 2 and is from LOTE into English.

  *N.B. Dialogue and Sight Translation are tested in one sitting. Subject areas are allocated to each student at random and will not be informed to the student beforehand.*

- **Two speeches interpreted in the consecutive mode**, of 300 words or equivalent length each, in two different subject areas, one from English into LOTE, and the other from LOTE into English, to be done in one sitting. Each 300-word speech will be divided into two segments of approximately 150 words per segment, and each of these 150-word segments will be interpreted one at a time. Subject areas of the two speeches will be informed to the students two week before the exam for preparation/research purposes.

  The content of the exam may not be the same as what was practised in class and may include complex language which an educated speaker of the language pair tested would be expected to be familiar with. Apart from the third component—consecutive interpreting, which the topic areas will be announced to students 2 weeks prior to the examination date as per NAATI regulations, topics for the other two components are chosen randomly and students cannot choose the topics in their testing.

  To be recommended to NAATI for accreditation, students must be marked by two examiners and achieve an average minimum mark of 70% in each
one of the three components: 1) Dialogues; 2) Sight translation; 3) Consecutive. Please refer Appendix 2 for NAATI marking sheet and RMIT marking descriptors.

For students who have achieved a minimum mark of 70% in the dialogue component but not in the other two or either of the two components, they will be recommended to NAATI for a conceded Paraprofessional Interpreter Accreditation.

**SECTION 3 ACCREDITATION EXAMINATION PROTOCOLS AND PROCEDURES**

**3.1 Translation Accreditation Examination**

Translation accreditation exams are conducted in a computer lab. Students must make sure during the semester that they have access to the necessary LOTE language on the computers in the lab. The use of published monolingual and bilingual dictionaries, glossaries and thesauri **IS** allowed. Students are advised that they should supply their own and are **NOT** allowed to share with other students.

Electronic Dictionaries (monolingual and bilingual) **ARE** allowed. However if the machine has sentence/phrase translation functions, they must not be used. Students may be asked to record the brand, model and type on their examination booklets.

**NB:** *Electronic dictionaries must be kept in mute mode at all times during the examination in order not to disturb other students; otherwise they will be confiscated and returned at the end of that exam session.*

Other books, reference materials or notes **are NOT allowed** into the examination venue. Mobile phones, smart phones, handwritten materials, class notes, newspaper and periodical articles and other reference materials **are NOT allowed**.

**3.2 Interpreting Accreditation Examination**

**DIALOGUE INTERPRETATION AND SIGHT TRANSLATION**

i) Students are required to interpret two bilingual dialogues **and** two sight translation passages **in one sitting**.

ii) The two dialogues are in two different topic areas.
iii) The first sight translation passage relates to the topic of the first dialogue and is interpreted from English into the LOTE.

iv) The second sight translation passage relates to the topic of the second dialogue and is interpreted from the LOTE into English.

v) The dialogue and sight translation components are sequenced as follows during the onesitting:
   - The student does dialogue 1.
   - The student does sight translation passage 1, from English into the LOTE. A timed three-minute reading time is given before delivery.
   - The student does dialogue 2.
   - The student does sight translation passage 2, from the LOTE into English. A timed three-minute reading time is given before delivery.

vi) Dictionaries –either electronic or printed-, glossaries, notes or other reference materials are NOT allowed during the dialogue and sight translation examination.

vii) Students CAN make notes in the hard copy sight translation scripts provided.

viii) Students are allowed one repeat of any segment in the dialogue without penalty. Additional repeats may attract penalties in the marking.

CONSECUTIVE EXAMINATIONS

i) Students are required to interpret a 300 word speech from English into LOTE, and a 300 word speech from LOTE into English in the consecutive mode in one sitting. Subject areas of the two speeches will be announced two weeks before the exam.

ii) Students listen to each passage once. Each passage is given in two segments of approximately 150 words per segment. No repeats are allowed. Students interpret one segment at a time.

iii) Students may take notes as they listen to each segment. They are not allowed clarifications or any other questions. They have 15 seconds at the end of each 150 word segment to scan their notes. They commence their interpretation at the end of these 15 seconds.

iv) The recommended time to complete the delivery of interpreting for each segment is 3.5 minutes. Unduly slow delivery will attract penalties
in marking.

v) The Examination Supervisor for each examination session is responsible for ensuring that the examination is conducted according to NAATI protocols.

### 3.3 RESULTS

All semester-end results will be published through the myRMIT by the University’s stipulated date (check student diary).

Before the official result release day, students are **NOT** under any circumstances, to ask examiners and teachers about marking progress or any other assessment information regarding the accreditation examinations or end of semester examinations.

No results are final until the Examiners’ Board and/or Student Progress Committee have met, duly considered all student progress issues and moderated, vetted and approved results for entry on the RMIT database.

*After results have been formally released by the University, students wishing to view their examination performance should contact the Program Coordinator who will advise them on the process in place at the time and make any necessary arrangements.*

Information about results, please check via:
http://www.rmit.edu.au/browse/Current%20students/Admin%20essentials/Assessment%20and%20exams/Results/

Information about transcript, please check via:
http://www.rmit.edu.au/browse;ID=vn1jt0uzqy99

### SECTION 4 HOW TO GRADUATE AND APPLY FOR NAATI ACCREDITATION

### 4.1 GRADUATION

For information on how to graduate, students should visit the following web link:

http://www.rmit.edu.au/browse;ID=q4n48yczl2jaz

Eligible students will graduate in absentia by default, if you do not select a graduation option. Your testamur (graduation certificate) will be posted to you. If you wish to attend a graduation ceremony, you **must** apply to graduate and make your graduation choice online.
4. 2  **NAATI ACCREDITATION**

NAATI does not award accreditations automatically.

**The following steps apply:**

i) The School sends a complete set of results to NAATI as soon as practicable after the completion of any program.

ii) Graduates need to have a complete transcript of results, as issued by RMIT, with no pending or incomplete results.

iii) Student must ensure that they have achieved an accreditation, i.e. a result of DISTINCTION or above in the particular accreditation course.


v) Student should post the form to NAATI at the address shown in the form.

vi) After receiving your application, NAATI will cross-check the documentation against the results provided to them by the School to verify eligibility to be accredited and will inform the applicant in due course.

vii) Graduates are not to contact or send partial documentation to NAATI. NAATI will not process your application unless you have provided a certified copy of your completed transcript and NAATI has received a set of results from the School.

viii) Graduates must wait until NAATI advises them that they have been accredited before applying for any certificates, identification cards, stamps, etc.

4.3 **NAATI REVALIDATION**

NAATI has introduced a revalidation scheme in which accreditation is available for a period of three years, after which time a process of revalidation is needed based on activity in the T&I field and professional development.

Details are on the NAATI website at: [www.naati.com.au](http://www.naati.com.au)
If you require any further clarification please contact the Program Coordinator, Mr Sedat Mulayim on: 9925 2328 (reception), 9925 9920 (direct) or sedat.mulayim@rmit.edu.au

Or

The Program Admin Officer, Mr Piero Damiani on 9925 2328 (reception), 9925 2995 (direct) or translating.interpreting@rmit.edu.au
APPENDIX 1

**TRANSLATOR EXAMINATION RESULT SHEET FOR NAATI ACCREDITATION**

<table>
<thead>
<tr>
<th>Year: 201_</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Sem _</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Translation (Eng into LOTE)</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Translation (LOTE into English)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Student’s Name:**

**Student ID No.:**

<table>
<thead>
<tr>
<th>Passage 1</th>
<th>Passage 2</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>weighting</td>
<td>50%</td>
<td>50%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Examiner 1 / 2</th>
<th>Examiner 1</th>
<th>Examiner 2</th>
<th>Examiner 1</th>
<th>Examiner 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Out of 50</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

You must achieve

at least 32.5 / 50 for each passage; and 70 / 100 for the total of both passages
to be recommended to NAATI for Professional Translator accreditation

**Positive Aspects of Performance**

- Understanding of source text
- Rendition of target text
- Innovative translating skills
- Other: please specify

**Negative Aspects of Performance**

- Lack of proficiency in English
- Lack of proficiency in LOTE
- Lack of translating skills
- Other: please specify

**NAATI Performance Descriptors:**

1. **Accuracy**

   - (A) Significant omissions
   - (B) Significant mistranslations
   - (C) Significant unjustified insertions
   - (D) Failure to complete passage

2. **Resources of Language – Comprehension of Original**

   - (E) Misunderstanding vocabulary
   - (F) Misunderstanding grammatical features
   - (G) Misunderstanding sentence structures

3. **Resources of Language – Expression of Translation**

   - (H) Inappropriate / inexact word choices
   - (I) Grammatical errors
   - (J) Inappropriate / awkward sentence structures
   - (K) Non-idiomatic usage
   - (L) Spelling errors
   - (M) Errors of punctuation / capital letters

4. **Technique**

   - (N) Translated too literally in some segments
   - (O) Translated too freely
   - (P) Excessive and / or incorrect paraphrasing

---

Assessor 1 (please circle & date) | Assessor 2 (please circle & date) | Date
### RMIT Descriptors on Overall Translating Performance:

In each column the assessor marks the box that best reflects the candidate’s performance in that dimension. The assessor may also insert, circle, and/or cross out words in a description to make the evaluation more specific.

<table>
<thead>
<tr>
<th>Grade</th>
<th>Source Language Competency</th>
<th>Target Language Competency</th>
<th>Capacity to Produce Specialist Texts</th>
<th>Translation Techniques</th>
<th>Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>HD 80+</strong></td>
<td>The translation demonstrates a high level of accuracy, containing virtually no undue distortions, omissions, or additions.</td>
<td>There are few (if any) grammatical/syntactical errors or issues with spelling, accentuation/use of diacritics, and punctuation. The translation uses cohesive devices appropriately, facilitating a completely coherent reading.</td>
<td>The translation conforms fully to conventions of the relevant genre, uses appropriate register and specialist terminology in the target language for the specified audience.</td>
<td>The translation demonstrates a high level of ability to employ translation techniques (eg, literal, free, dynamic, paraphrasing, loan words) to solve transfer issues and shows excellent awareness of professional practices of revising own work to finalise a translation.</td>
<td>The overall performance displays the candidate’s excellent ability to produce translation of specialist texts confidently.</td>
</tr>
<tr>
<td><strong>DI * 70-79</strong></td>
<td>The translation demonstrates a good level of accuracy, containing sporadic undue distortions, omissions, or additions.</td>
<td>There are a small number of grammatical/syntactical errors or issues with spelling, accentuation/use of diacritics, and punctuation. The translation uses cohesive devices and facilitates a coherent reading.</td>
<td>The translation conforms to conventions of the relevant genre, uses appropriate register and specialist terminology with minor infelicities in the target language for the specified audience.</td>
<td>The translation demonstrates a good level of ability to employ translation techniques (eg, literal, free, dynamic, paraphrasing, loan words) to solve transfer issues and shows awareness of professional practices of revising own work to finalise a translation.</td>
<td>The overall performance displays the candidate’s ability to produce translation of specialist texts confidently.</td>
</tr>
<tr>
<td><strong>CR 60-69</strong></td>
<td>The translation demonstrates a fair level of accuracy, containing some undue distortions, omissions, or additions.</td>
<td>There are a some grammatical/syntactical errors or issues with spelling, accentuation/use of diacritics, and punctuation. The use of cohesive devices is sometimes inaccurate which slightly hampers a coherent reading.</td>
<td>The translation by and large conforms to conventions of the relevant genre, uses appropriate register and specialist terminology with some infelicities in the target language for the specified audience.</td>
<td>The translation demonstrates some level of ability to employ translation techniques (eg, literal, free, dynamic, paraphrasing, loan words) to solve transfer issues and shows limited awareness of professional practices of revising own work to finalise a translation.</td>
<td>The overall performance displays the candidate’s ability to produce translation of specialist texts which need some rework.</td>
</tr>
<tr>
<td><strong>PASS 50-59</strong></td>
<td>The translation demonstrates a low level of accuracy, containing many undue distortions, omissions, or additions.</td>
<td>There are a many grammatical/syntactical errors or issues with spelling, accentuation/use of diacritics, and punctuation. The use of cohesive devices is inaccurate which hampers a coherent reading.</td>
<td>The translation does not completely conform to conventions of the relevant genre, uses somewhat inappropriate register and specialist terminology with many infelicities in the target language for the specified audience.</td>
<td>The translation demonstrates a low level of ability to employ translation techniques (eg, literal, free, dynamic, paraphrasing, loan words) to solve transfer issues and shows low awareness of professional practices of revising own work to finalise a translation.</td>
<td>The overall performance displays the candidate’s ability to produce translation of specialist texts which need major rework.</td>
</tr>
<tr>
<td><strong>FAIL Under 49</strong></td>
<td>The translation demonstrates an unacceptably low level of accuracy, containing excessive undue distortions, omissions, or additions.</td>
<td>There are a excessive grammatical/syntactical errors or issues with spelling, accentuation/use of diacritics, and punctuation. The use of cohesive devices is grossly inaccurate which completely hampers a coherent reading.</td>
<td>The translation does not conform to conventions of the relevant genre, uses completely inappropriate register and specialist terminology with excessive infelicities in the target language for the specified audience.</td>
<td>The translation demonstrates almost no ability to employ translation techniques (eg, literal, free, dynamic, paraphrasing, loan words) to solve transfer issues and shows almost no awareness of professional practices of revising own work to finalise a translation.</td>
<td>The overall performance displays the candidate’s ability to produce translation of specialist texts which need complete rework.</td>
</tr>
<tr>
<td><strong>DNS</strong></td>
<td>Did Not Submit for Assessment</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---

**Descriptors**

- **HD 80+** (High Distinction): Demonstrates a high level of accuracy and coherence, few if any grammatical or syntactical errors. Translation uses cohesive devices appropriately.
- **DI 70-79** (Distinction): Demonstrates a good level of accuracy, containing sporadic undue distortions. Translation uses cohesive devices and facilitates coherence.
- **CR 60-69** (Credit): Demonstrates a fair level of accuracy. Translation contains some undue distortions, but cohesive devices are accurate enough to maintain coherence.
- **PASS 50-59** (Pass): Demonstrates a low level of accuracy. Translation contains many undue distortions, affecting the coherence.
- **FAIL Under 49** (Fail): Demonstrates an unacceptably low level of accuracy. Translation contains excessive undue distortions, affecting the overall coherence.
- **DNS** (Did Not Submit for Assessment): No evaluation is possible as the candidate did not submit their work.
## APPENDIX 2

### INTERPRETING EXAMINATION RESULT SHEET
FOR NAATI ACCREDITATION

<table>
<thead>
<tr>
<th>Language:</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Student’s Name:</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Student ID No.:</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Dialogue 1</th>
<th>Dialogue 2</th>
<th>Sight Trans Eng&gt;LOTE</th>
<th>Sight Trans LOTE&gt;Eng</th>
<th>Monologue Eng&gt;LOTE</th>
<th>Monologue LOTE &gt; Eng</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weighting</td>
<td>25%</td>
<td>25%</td>
<td>10%</td>
<td>10%</td>
<td>15%</td>
<td>15%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Marks out of 100</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
</table>

Average of Each Component

---

You must achieve an average of at least 70% (out of 100) in each component in order to be recommended to NAATI for Professional Interpreter accreditation.

A conceded Paraprofessional Interpreter accreditation may be awarded when only the average of the dialogue component achieves 70% or above.

### Aspects of positive performance
- □ High level of English proficiency
- □ High level of LOTE proficiency
- □ Demonstration of transferring skills
- □ Other: please specify

### Aspects of negative performance:
- □ Lack of proficiency in English
- □ Lack of proficiency in LOTE
- □ Lack of interpreting skills
- □ Other: please specify

### NAATI Marking Descriptors:

#### 1. Accuracy
- (A) □ Significant omissions
- (B) □ Significant distortions
- (C) □ Significant unjustified insertions
- (D) □ Excessive requests for repetition
- (E) □ Candidate abandoned test

#### 2. Language Proficiency – LOTE > Eng
- (F) □ Inadequate comprehension of LOTE
- (G) □ Inappropriate word choices in English
- (H) □ Incorrect sentence structures in English
- (I) □ Grammatical errors in LOTE
- (J) □ Unidiomatic usage in English
- (K) □ Unsatisfactory pronunciation in English

#### 3. Language Proficiency – Eng > LOTE
- (L) □ Inadequate comprehension of English
- (M) □ Inappropriate word choices in LOTE
- (N) □ Incorrect sentence structures in LOTE
- (O) □ Grammatical errors in LOTE
- (P) □ Unidiomatic usage in LOTE
- (Q) □ Unsatisfactory pronunciation in LOTE

#### 4. Technique
- (R) □ Excessive pauses / hesitations
- (S) □ Note-taking interfered with flow of interpreting
- (T) □ Inappropriate strategy when seeking clarification/repetition
- (U) □ Excessive number of clarifications/repetitions

#### 5. Other
- (V) □ Inadequate comprehension of scenarios

---

<table>
<thead>
<tr>
<th>Assessor 1 (please circle &amp; date)</th>
<th>Assessor 2 (please circle &amp; date)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Date
### RMIT Descriptors for Overall Interpreting Performance:

In each column the assessor marks the box that best reflects the candidate’s performance in that dimension. The assessor may also insert, circle, and/or cross out words in a description to make the evaluation more specific.

<table>
<thead>
<tr>
<th>Grade</th>
<th>Accuracy</th>
<th>Repet'n/Clarific'n</th>
<th>Qual of Lang</th>
<th>Techniques</th>
<th>Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>CHD 80+</strong></td>
<td>The interpreting contains virtually no undue distortions, omissions, or additions.</td>
<td>There are virtually no undue requests for repetitions/clarifications.</td>
<td>There are few (if any) grammatical errors, issues with register, collocations and pronunciation, and these are not serious enough to impede comprehension or to cause distortion.</td>
<td>The techniques for managing discourse, overall presentation and fluency are at an acceptable industry level.</td>
<td>The overall performance displays the candidate’s excellent ability to handle interpreting in complex settings confidently.</td>
</tr>
<tr>
<td><strong>CDI * 70-79</strong></td>
<td>The interpreting contains some minor distortions, omissions, or additions, which are mainly at the word/phrase level.</td>
<td>There are a few requests for repetitions/clarifications but these are either justified or do not disrupt the flow of communication.</td>
<td>There are some grammatical errors, issues with register, collocations and pronunciation, but not serious enough to impede comprehension or to cause distortion.</td>
<td>The techniques for managing discourse, overall presentation and fluency are at an acceptable industry level, with only minor problems.</td>
<td>The overall performance displays the candidate’s good ability to handle interpreting in complex settings confidently.</td>
</tr>
<tr>
<td><strong>CC 60-69</strong></td>
<td>The interpreting contains a few minor and major distortions, omissions and additions, some of which are beyond the word/phrase level.</td>
<td>There are a few requests for repetitions/clarifications but many of them are not justified and/or some impede the flow of communication.</td>
<td>There are some grammatical errors, issues with register, collocations and pronunciation, and some are serious enough to impede comprehension or to cause distortion to miscommunication.</td>
<td>The techniques for managing discourse, overall presentation and fluency are below acceptable industry level, with some major problems.</td>
<td>The overall performance is below the acceptable industry standards for handling interpreting in complex settings.</td>
</tr>
<tr>
<td><strong>PASS 50-59</strong></td>
<td>The interpreting contains many minor and some major distortions, omissions and additions, some of which are beyond the word/phrase level.</td>
<td>There are many requests for repetitions/clarifications but many of them are not justified and/or they impede the flow of communication.</td>
<td>There are many grammatical errors, issues with register, collocations and pronunciation, and some are serious enough to impede comprehension or to cause distortion to miscommunication.</td>
<td>The techniques for managing discourse, overall presentation and fluency are below acceptable industry level, with many major problems.</td>
<td>The overall performance is well below the acceptable industry standards for handling interpreting in complex settings.</td>
</tr>
<tr>
<td><strong>FAIL Under 49</strong></td>
<td>The interpreting contains many minor and many major distortions, omissions and additions which have contributed to a breakdown of communication.</td>
<td>There are too many requests for repetitions/clarifications which impede the flow of communication.</td>
<td>There are too many grammatical errors, issues with register, collocations and pronunciation, and many are serious enough to impede comprehension or to cause distortion to miscommunication.</td>
<td>The techniques for managing discourse, overall presentation and fluency are well below acceptable industry level, with too many major problems.</td>
<td>The overall performance is a long way below the acceptable industry standards for handling interpreting in complex settings.</td>
</tr>
<tr>
<td><strong>DNS</strong></td>
<td>Did Not Submit for Assessment</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(* MINIMUM STANDARD REQUIRED FOR NAATIACCREDITATION)
## Master of Translating and Interpreting
And

**Graduate Diploma in Translating and Interpreting**

### Section 1 - Student Information

<table>
<thead>
<tr>
<th>Title</th>
<th>Mr → Ms → Mrs → Other → (specify)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Family Name</td>
<td></td>
</tr>
<tr>
<td>Given Name/s</td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td>Male → Female →</td>
</tr>
<tr>
<td>Telephone (BH)</td>
<td></td>
</tr>
<tr>
<td>Mobile Telephone</td>
<td></td>
</tr>
<tr>
<td>Language Group</td>
<td></td>
</tr>
<tr>
<td>Enrolled</td>
<td>Part-time → Full-time →</td>
</tr>
<tr>
<td>Student IDNumber</td>
<td></td>
</tr>
</tbody>
</table>

### Section 2 - Declaration by Student

- I have received and read the above information which I agree to retain for future reference and comply with accordingly.

- I understand and agree that the program will disclose my result in the accreditation course(s) to NAATI for verification purposes.

Student's Signature: …………………………………………… Date: ……….../………....

### Section 3 – Affix one passport photo for internal program office use only.

---

*This acknowledgement form must be returned to Program Coordinator via L.2 B.37 Reception before Census Day.*