Function-based Translation Quality Assessment

Rudy Sofyan[1], Bahagia Tarigan[1]

[1] *Linguistics Department, Universitas Sumatera Utara, Jalan Abdul Hakim Kompleks USU, Medan, Indonesia*

Keywords: Translation Quality Assessment, TQA Model, Translation Function.

Abstract: This article is aimed at (i) finding out the results of TQA using three TQA models, and (ii) developing a TQA model. The research used a descriptive qualitative method using a process-oriented translation as the approach. The data were the results of the assessment and the questionnaires from five raters. The raters were translation experts and professional translators who were asked to assess the TT by using three different TQA models. The first model (Model A) was proposed by Hurtado (1995), the second model (Model B) was proposed by Waddington (2001), and the third model (Model C) was proposed by Nababan et al. (2012). The TT assessed was entitled "*Sejarah Awal Yellowstone National Park*" translated by a professional translator. This research found that (i) using the three different TQA models results in different level of quality of the TT caused by two factors: the absence of text/sentence function as the quality aspect and the limited description of quality level; (ii) the model developed in this article is called function-based translation quality assessment using a holistic method of assessment in which the whole criteria of translation quality (accuracy, finding equivalents, translation skill, text/sentence function and grammar and TT style) are assessed.

1 INTRODUCTION

Quality translation is an absolute condition to be achieved by every translator because of which, as Williams (2004) argues, translation quality has become a central issue in a product-oriented translation studies for such factors as aesthetic, religious, political, pedagogical, administrative and economic. Nevertheless, such factors generate various questions concerning the criteria of quality translation product. In addition, the definition of translation quality as the translation producing a good target text (TT), proposed by Schäffner (1997), certainly brings about more various questions such as "What is a good TT?", "What are the characteristics of a good TT?", or "What are the elements of the source text (ST) that need to be added, substituted or omitted in the TT to produce a good TT". Therefore, it is necessary to design and develop a model that hopefully would be able to distinguish different level of translation quality.

To date, translation quality assessment (TQA) continues to receive serious attention from the researchers in translation studies. Studies on the theory of translation assessment (Farahzad, 1992; Hurtado, 1995; House, 1998; McAlester, 2000; Bowker, 2001; Melis and Hurtado, 2001; Williams,

2001, 2009; Waddington, 2001; Khanmohammad and Osanloo, 2009; Nababan, Nuraeni, and Sumardino, 2012) produced several models of TQA. In his research on the quality of translations into a foreign language, McAlester (2000) states that the methods to be used in TQA should be reliable, valid, objective, and practical. In relation to McAlister's reliability, validity, objectivity and practicality in assessing translation quality, Williams (2004) proposes that TQA could be done by using either (1) models with a quantitative dimension and (2) models with non-quantitative dimension, text-logical models, or both. The second type refers to House's (1998) models of TQA. In his further study, Williams (2009) argues that TQA could be qualitative or quantitative where the assessment could be based either on mathematical or statistical measurements or on reader response, interviews and questionnaires.

Although many experts in translation studies have designed and developed various models of TQA, the problem of TQA does not stop because certain TT might have different quality when assessed by different models of TQA. This is a serious problem that needs serious attention and exploration to find out the factors leading to such problem. This research is aimed at (i) finding out the results of TQA using three models of TQA, and (ii) developing a model of TQA.

Table 1: Hurtado's error analysis in TQA.

No.	Evaluated Aspects	Minor Errors (-1)	Serious Errors (-2)
1	Providing improper equivalents affecting the ST understanding		
	a. Contresens	(-1)	(-2)
	b. Faux sens	(-1)	(-2)
	c. Nonsens	(-1)	(-2)
	d. Addition	(-1)	(-2)
	e. Omission	(-1)	(-2)
	f. Unresolved extralinguistic references	(-1)	(-2)
	g. Loss of meaning	(-1)	(-2)
	h. Inappropriate linguistic variation (register, style, dialect, etc.).	(-1)	(-2)
2	Providing improper expressions in the TT		
	a. Spelling	(-1)	(-2)
	b. Grammar	(-1)	(-2)
	c. Lexical items	(-1)	(-2)
	d. Text	(-1)	(-2)
	e. Style	(-1)	(-2)
3	Providing insufficient equivalents affecting the transfer of either the main function or secondary functions of the ST.	(-1)	(-2)
4	The plus point	Good Solutions (+1)	Exceptionally Good Solutions (+2)

The first model of TQA is developed by Hurtado (1995). The model is developed mainly based on error analysis that classifies the error into either serious error penalized with (-2) points or minor error penalized with (-1) point. The errors are grouped into three categories as shown in Table 1.

In addition to the three categories of errors, this model also provides additional category containing the plus points: good solutions to translation problems are awarded +1 point and exceptionally good ones are awarded +2 points.

Unlike the first model emphasizing on error analysis, the second model developed by Waddington (2001) adopts a holistic method of assessment. Waddington (2001) designs a five-level scale of assessment whose description of translation quality is shown in Table 2.

Table 2: Waddington's holistic model of TQA.

Level		Quality of expression in TL	Degree of task completion	Mark
Level 5	The ST content is completely transferred. Only minor revision is needed to reach professional standard.	Almost all the translation reads like a text originally written in English. There may be minor lexical, grammatical or spelling errors.	Successful	9, 10
Level 4	The ST content is almost completely transferred. There may be one or two insignificant inaccuracies. The TT requires certain amount of revision to reach professional standard.	Most of the translation reads like a text originally written in English. There are a number of lexical, grammatical or spelling errors.	Almost completely successful	7, 8
Level 3	The transfer of the general idea(s) contains a number inaccuracies. The TT needs considerable revision to reach professional standard.	Certain parts of the translation read like a text originally written in English, but others read like a translation. There are a considerable number of lexical, grammatical or spelling errors.	Adequate	5, 6
Level 2	The transfer of the general idea(s) contains serious inaccuracies. The TT needs a lot of revision to reach professional standard.	Almost the entire text reads like a translation; there are continual lexical, grammatical or spelling errors.	Inadequate	3, 4
Level 1	The transfer of the ST content is totally inadequate. The translation is not worth revising.	The candidate reveals a total lack of ability to express him/herself adequately in English.	Totally inadequate	1, 2

The third method is developed by Nababan et al. (2012). Like Waddington's model, this is also a holistic model which assesses the translation quality through three instruments, i.e. accuracy, acceptability and readability. Each of the instruments is divided into three parts: translation categories, scores and qualitative parameters. The score ranges from 1 (describing inaccuracy, unacceptability and low readability level) to 3 (describing accuracy, acceptability and high readability level). Even though the model uses three different instruments, it is a holistic assessment because the scores from each instrument will be combined to produce the final score which eventually describes the quality of the assessed translation (see Table 3).

Table 3: Nababan's holistic model in TQA.

No.	Quality Aspects Evaluated	Mark
1	Accuracy	n x 3
2	Acceptability	n x 2
3	Readability	n x 1

Letter 'n' in Table 3 represents the mark given for the aspect of accuracy, acceptability and readability. The highest score (x3) given to accuracy aspect shows that Nababan's model gives more emphasis on accuracy in assessing the overall quality of the translation. The final mark of the assessment is obtained by dividing the sum of the quality aspects evaluated with 6.

2 RESEARCH METHODOLOGY

This research used a descriptive qualitative method using a process-oriented translation as the approach. The data were the results of the assessment and the questionnaires from five raters. The raters were translation experts and professional translators who were asked to assess the TT by using three different models of TQA. The first model (Model A) was proposed by Hurtado (1995), the second model (Model B) was proposed by Waddington (2001), and the third model (Model C) was proposed by Nababan et al. (2012). The TT that was assessed was entitled "*Sejarah Awal Yellowstone National Park*" which was translated by a professional translator from English whose title was "Early History of Yellowstone National Park". In the process of data collection, before the raters assessed the TT, they were given the explanation of how the three models of TQA work. Having assessed the TT, they were given an open questionnaire containing their experience in applying those models. Then the results

of the assessment and questionnaires were analyzed to find out whether their applications resulted in similar level of the TT quality. Then, based on this analysis, a new model of TQA was developed.

3 RESULTS AND DISCUSSION

The data analysis shows that assessing the translation by using the three different models of TQA applied by the five raters results in different level of quality of the TT. Table 4 shows the results of assessment using Model A.

Table 4: The results of TQA using model A.

No.	Raters	Translation Quality	
		Score	Description
1	Rater A	82	Good
2	Rater B	81	Good
3	Rater C	81	Good
4	Rater D	78	Fairly Good
5	Rater E	83	Good

The results of the assessment from the five raters using Model a show the final scores of 82, 81, 81, 78 and 83, respectively. This means that four out of five raters agree that the TT quality is good, and only doe's one rater state that the TT quality is fairly good.

Meanwhile, the results of the assessment from the five raters using Model B (see Table 5) show the final scores of 70, 61, 68, 60 and 61, respectively. This indicates that four out of five raters agree that the TT is almost completely successful.

Table 5: The results of TQA using model B.

No.	Raters	Translation Quality	
		Score	Description
1	Rater A	70	Almost completely successful
2	Rater B	61	Almost completely successful
3	Rater C	68	Almost completely successful
4	Rater D	60	Adequate
5	Rater E	61	Almost completely successful

The results of TQA presented in Table 5 show that none of the raters evaluates the TT as a successful translation product, or a TT with a good quality. This is reflected through the different score of the TT displayed in Table 4 and Table 5.

Furthermore, by using the scoring system of Nababan et al. (2012), four out of five raters agree that

the quality of the TT is average. Their scores are 2.50, 2.17, 2.33, 2.17, 2.33, respectively, as displayed in Table 6.

Table 6: The results of TQA using model C.

No.	Raters	Translation Quality	
		Score	Description
1	Rater A	2.50	Good
2	Rater B	2.17	Average
3	Rater C	2.33	Average
4	Rater D	2.17	Average
5	Rater E	2.33	Average

These results lead to the finding that the TT has different level of quality when assessed by using different models of TQA. Although it is impossible to get the exactly similar results of assessment by using different models or methods of TQA, the range should not be too wide because it will result in the assessment uncertainty in terms of which translation has good quality. Consequently, the finding needs to be further explored since any models used in assessing translation should arrive at the same or nearly the same level of quality. The main cause of such different assessment results is the absence of evaluation on the text or clause function in the parameters or instruments used in the three models. Look at the translation in (1).

(1) ST : Native Americans have first claim on the Yellowstone Plateau and lived in the area in peaceful tranquility until the early 1800s--undisturbed by the presence of white men.

TT : *Penduduk asli Amerika pertama sekali mendiami Yellowstone Plateau dan tinggal di daerah itu dengan damai sampai awal tahun 1800an--tanpa diganggu oleh kehadiran orang kulit putih.*

Using Model A, Rater A found two serious errors in the form of contresens (mistranslation) and gave - 4 points for the sentence in (1). The first contresens was keeping ST phrase "Yellowstone Plateau" in the TT. Keeping the ST "Yellowstone Plateau" in the TT belongs to mistranslation because in bahasa Indonesia has the equivalent of "Plateau", i.e. "*Dataran Tinggi*". Although keeping the ST word(s) is acceptable in translation, it is not proper to be used in (1) because it might make the TT readers unable to understand what the "Yellowstone Plateau" is exactly (whether it is a highland, lowland, rocky area, etc.). Meanwhile, the second contresens was writing "*pertama sekali mendiami*" as the equivalent of the ST "have first claim". The mistranslation occurred

when the translator decided to omit or substitute the meaning of "claim" which, according to the raters, played a very important role in this sentence. Omitting the meaning of "claim" shows the failure in providing the right equivalent. Nevertheless, Rater A also gave +2 points for good solutions in translating 2 ST phrases. The good solutions found in the TT were writing "*tinggal di daerah itu dengan damai*" as the equivalent of the ST phrase "lived in the area in peaceful tranquility" and "*tanpa diganggu oleh kehadiran orang kulit putih*' as the equivalent of the ST phrase "undisturbed by the presence of white men". Based on these minus and plus points, the assessment for the sentence in (1) is -2 points.

Meanwhile, following Model B, Rater A found almost complete meaning transfer due to several insignificant inaccuracies requiring revisions to reach professional standard. Such inaccuracies are due to the use of literal translation technique, incorrect equivalents and borrowing. Literally translating the ST word "first" into the TT phrase "*pertama sekali*" shows inaccuracy as the whole meaning of the sentence in (1) has nothing to do with sequence of actions. Similarly, the TT word "*mendiami*" is the incorrect equivalent of the ST word "claim" because the ST term "first claim" is commonly equivalent with the TT term "*memiliki hak*". Moreover, borrowing the ST word "Plateau" is inaccurate because it has its established equivalent in bahasa Indonesia "*Dataran Tinggi*". Due to such inaccuracies, Rater A gave a score 61, and based on the degree of task completion, it was viewed as "almost completely successful" translation.

Furthermore, following Model C, Rater A found meaning distortion that bothers the complete transfer of meaning from the ST to the TT. Based on the accuracy instrument, as there are only three levels of accuracy (accurate, less accurate and inaccurate), Rater A classified the sentence in (1) as less accurate translation. Based on the data from the questionnaire, the quality of sentence in (1) could have been more than "less accurate" because it is between "accurate" and "less accurate". This is one of the weaknesses of scoring system dividing quality only into three levels.

The results of assessment performed by Rater A highlighted in the previous paragraphs represent the results of assessments from the other four raters that also show different level of quality of the TT when assessed by using different models of TQA. Based on the results of the questionnaire, Rater B reports that the difference in grading the quality of the TT is caused by the absence of detailed assessment criteria. He says that Waddington's holistic model is very good because it assess the whole quality of the TT;

however, its description for each level of quality is not as detailed as the one mentioned in Hurtado's error analysis model. For example, the quality description "Almost all the translation reads like a piece originally written in English" does not provide a clear description of which quality that resembles or does not resemble the ST. In addition, the absence of assessing the text and sentence function is also the problem that makes different models of TQA assess the same TT in different quality.

The big differences found in the assessment results can be resolved when the three models provide one more quality aspect to be evaluated, i.e. text or sentence function. By having this aspect, which certainly has its own role in the scoring system, such big differences can be minimized because the scoring of the other quality aspects also needs to be adjusted. In addition to the additional aspect of translation quality (i.e. text/sentence function), the three models also need more comprehensive scoring system distinguishing the quality level within the same aspect of quality. For example, when the highest score for a certain quality aspect is 30, the description of how such highest score is obtained should be described, and the description should be further continued to the other lower scores.

Based on these findings, this article proposes a model of TQA called "Function-based translation quality assessment". This model is also a holistic model since it evaluates or assesses the whole translation quality criteria in the TT. Suggested by its name, this model includes text or sentence function as one of the criteria of TQA. This supports the idea that the quality of the TT should not be assessed only based on the elements constructing the text, but it should also evaluates whether the TT has achieved the function intended by the ST (Sofyan and Tarigan, 2018). In addition to scoring system and criteria of quality proposed by Hurtado (1995), Waddington (2001) and Nababan et al. (2012), this model provides more detailed quality description, as suggested by Farahzad (1992) and Khanmohammad and Osanloo (2009), together with its more detailed scoring system (Sofyan, 2016). In general, this model divides the quality criteria into five broad categories: (i) accuracy, (ii) finding equivalents, (iii) translation skill, (iv) text/sentence function, and (v) grammar and TT style.

Following Nababan et al. (2012), accuracy is given the highest percentage (30%) of the total score (100) with the details described in Table 7.

Table 7: Description of accuracy.

Score Range	Description
25-30	There are no identifiable problems of comprehension to the TL readers. The original message is completely transferred to the TT without omissions or additions.
21-24	The problems of comprehension are only related to the most highly specialized vocabulary that do not affect the TL readers' understanding. There are some partial omissions and additions.
16-20	There is some difficulty experienced by the TL readers in understanding the TT due to the translator's misunderstanding of some parts of original message, apparent omissions and additions.
11-15	The ideas are poorly expressed. The translator's serious problems in understanding the ST disturb the transfer of the original message. The TT is difficult to understand.
1-10	The translator's serious problems in understanding the ST terribly disturb the transfer of the original message. The TL readers are unable to understand what the original writer intends to convey.

As accuracy determines 30 percent of the TT quality, the highest score for accuracy is 30. Each level of accuracy is given its own quality description which clearly distinguishes itself from other higher or lower level of accuracy. Besides, different quality descriptions are given different score range. By using such quality description, the quality criteria are more comprehensible and the assessment on the accuracy will be more representative.

Following Hurtado (1995), Waddington (2001) and Khanmohammad and Osanloo (2009), finding equivalent is considered as the second most important criteria of assessing the quality of TT, and in this model, it determines 25 percent of translation quality whose detailed descriptions are presented in Table 8.

Table 8: Description of finding equivalents.

Score Range	Description
20-25	All lexical and syntactic elements are understood through precise vocabulary usage. The words are chosen so skillfully that the translation reads like a good publishable text.
15-19	All lexical and syntactic elements are generally understood through good usage of a wide range of vocabulary and structures.

	Specialized vocabulary presents some problems with inappropriate equivalents.
10-14	Most of lexical and syntactic elements are generally understood through a fair range of vocabulary despite some observable gaps. Some vocabulary is inappropriately used. There is some evidence revealing translator's difficulties in finding equivalents, perception, wordplay and other linguistic features.
5-9	Comprehension of vocabulary and structures shows quite observable gaps which blur sense. There are problems in finding correct vocabularies due to translator's inability to overcome specialized vocabulary problems.
1-4	The vocabulary is inapprpriately used. Translator's poor comprehension of the ST seriously hinders the process of finding equivalents. As a whole, translation makes little sense.

The next quality is evaluated based on the translation skill criteria described in Table 9. Translation skill is evaluated based on the suggestions from Hurtado (1995) who considers that good translation skill should be given plus points to complete his error analysis model of TQA. He divides translation skill into two major skills: good solutions and exceptionally good solutions to translation problems. In addition, the description of translation skill criteria presented in this model also adapts the model proposed by Farahzad (1992), Khanmohammad and Osanloo (2009) and Sofyan (2016).

Table 9: Description of translation skill.

Score Range	Description
17-20	The translator demonstrates able and creative solutions to translation problems and is skillful in using resource materials.
13-16	The translator demonstrates consistent ability in identifying and overcoming translation problems. The TT contains only very few minor errors. There are no obvious errors in using resource materials.
9-12	The translator demonstrates a general ability to identify and overcome translation problems. There is a major translation error and/or an accumulation of minor errors. The improper use of reference materials is possibly reflected in the TT.
5-8	The translator demonstrates some difficulty in identifying and/or overcoming translation problems. There are several major translation errors and/or a large number of minor errors.

	The improper use of reference materials is reflected in the TT.
1-4	The TT reflects the translator's inability to identify and overcome common translation problems. There are numerous major and minor translation errors. The TT is seriously flawed translation. The TT contains improper use of materials and resources.

As described in Table 9, translation skill determines 20 percent of the total score of the TT quality; in other words, the highest score for translation skill is 20. By using the score range provided for each quality description, the TT assessment in terms of translation skill is expected to be more representative and reliable.

The next quality criteria developed in this model are related to text/sentence function, i.e. to assess whether the TT has met the text or sentence function intended in the ST (see Table 10). This is the main difference of this model from other previous models of TQA. The inclusion of sentence function is to show that this model, in addition to its nature as a holistic model, also assesses the quality at the sentence level (the function of the sentences that construct the TT). At the text level, the assessment is done to evaluate whether the TT is corresponding to the ST function (informative, expressive or operative) (Reiss, 1977;1989). At the sentence level, the evaluation is whether the sentence is used to inform, retell, report, describe, persuade or discuss.

Table 10: Description of text/sentence function.

Score Range	Description
13-15	Meeting the text/sentence function of the ST; creative inventions and skillful solutions to achieve the function of the ST; corresponding to the text function based on the TL perspective
10-12	Almost meeting the text/sentence function of the ST; some inventions to achieve the function of the ST; corresponding to the text function based on the TL perspective
7-9	Inconsistency in meeting the text/sentence function of the ST; awkward structure in achieving the ST function; not fully corresponding to the text function based on the TL perspective
4-6	Less attention to the text/sentence function of the ST; not corresponding to the text function based on the TL perspective
1-3	In contrast to the text/sentence function of the ST; not corresponding to the text function based on the TL perspective

As described in Table 10, text/sentence function determines 15 percent of the total score of the TT quality; in other words, the highest score for text/sentence function is 15. This means that if the TT should reach the highest level of quality, it must meet the text function of the ST corresponding to the TL perspective of text function.

The last quality criteria described in this model are related to the use of good grammar and TT style. Following Waddington's (2001) suggestions considering language (grammar and style) error as a minor error, the highest score for good language performance in translation is 10, meaning that grammar and style determine only 10 percent of the total score. The detailed description of each level of quality in grammar and style is presented in Table 11.

Table 11: Description of grammar and TT style.

Score Range	Description
9-10	The TT needs no improvement in terms of its grammatical and stylistic points though one or two natural failings are possibly observable. The TT shows native-like fluency in grammar.
7-8	The TT needs a little improvement in terms of its grammatical and stylistic points. The TT maintains advanced proficiency in grammar; There are some grammatical problems, but they do not affect the message.
5-6	The TT tends to have awkward grammatical usage due to literal rendering of the ST message, but it does not significantly hinders the message transfer. There are some attempts to reflect stylistic features of the TL. Some grammatical problems are apparent and have negative effects on message transfer.
3-4	The TT is awkward due to nonsensical grammatical usages. The TT sounds unnatural. There is a little attempt to reflect stylistic features of the TL. There is evidence of clear difficulties in following the TL style. Grammatical review of some areas is obviously required.
1-2	Little sense of the TL style. Knowledge of TL grammar is inadequate. The use of TL grammar is inadequate. Serious grammatical problems truly affects the message transfer.

As shown in Table 11, the highest score in grammar and TT style (10) will be obtained if the sentences in the TT are grammatically correct and corresponds to the TT style. The other lower levels of quality have limitations and are graded based on the number of limitations contained in the TT. The complete model of TQA developed in this article, function-based translation quality assessment, can be seen in the appendix.

4 CONCLUSIONS AND SUGGESTIONS

The findings elaborated in this article show that using different models of TQA in assessing the TT results in different level of quality. The TT has a good quality when assessed by Model A and Model C, but it has lower level of quality when assessed by using Model B. These findings need to be further explored to find the factors leading to such differences as using any model of TQA in assessing the TT should have arrived at the same or nearly the same level of translation quality. The results of analysis found that the absence of the quality aspect in the form of text/sentence function and the limited description of quality level are the main causes leading to such different assessment results. These findings, therefore, provide a chance to develop a model that accommodates the absence of the two main factors.

The function-based translation quality assessment uses a holistic method of assessment in which the whole criteria of translation quality are assessed. This model assesses the quality of translation under five aspects with different weight depending on how big they contribute to a good quality of translation. The five aspects are: accuracy (30%), finding equivalents (25%), translation skill (20%), text/sentence function (15%) and grammar and TT style (10%). Each of the aspects is provided with detailed descriptions to make the results of assessment more representative.

This model of TQA has not yet been tried out; therefore, it is suggested for other researchers to conduct a study on TQA by applying this model. The results of the studies applying this model are expected to give contributions in the improvement of this model.

ACKNOWLEDGEMENTS

REFERENCES

Bowker, L. 2001. Towards a methodology for a corpus-based. Approach to translation evaluation. *Meta*, *46*(2), 345–64.

Farahzad, F. 1992. Testing achievement in translation classes. In Dollerup, C. & Loddergard, A. (Eds.), *Teaching translation and interpreting*. Benjamins, Amsterdam/ Philadelphia. pp. 271-278.

House, J. 1998. Quality of translation, In Baker, M. (Ed.), *Routledge Encyclopedia of Translation Studies*. Routledge, London.

Hurtado, A. A. 1995. La didáctica de la traducción. Evolución y estado actual. In Fernández, P. (Ed.) *X Perspectivas de la Traducción*. Universidad de Valladolid, Valladolid. pp.49-74.

Khanmohammad, H. & Osanloo, M. 2009. Moving toward objective scoring: A rubric for translation assessment. *JELS*, *1*(1), 131-53.

McAlester, G. 2000. The evaluation of translation into a foreign language. In Schaeffner, C. & Adab, B. (Eds.), *Developing translation competence*. Benjamins Translation Library, 38. Benjamins, Amsterdam. pp. 229-42.

Melis, N. M. & Hurtado, A. A. 2001. Assessment in translation studies: Research needs. *Meta*, *46*(2), 272–87.

Nababan, M., Nuraeni, A. & Sumardiono. 2012. Pengembangan model penilaian kualitas terjemahan. *Kajian Linguistik dan Sastra*, *24*(1), 39-57.

Reiss, K. 1977/1989. Text-types, translation types and translation assessment (Translated by A. Chesterman). In Chesterman, A. (Ed.), *Readings in translation theory*. Oy Finn Lectura Ab, Helsinki. pp. 105-15.

Schäffner, C. 1997. From 'good' to 'functionally appropriate': Assessing translation quality. *Current Issues in Language and Society*, *4*(1), 1-5.

Sofyan, R. 2016. Translation process and translation quality: A study of Indonesian student translators. *Doctoral Dissertation*. Universitas Sumatera Utara, Medan.

Sofyan, R. & Tarigan, B. 2018. Theme markedness in the translation of student translators. *Indonesian Journal of Applied Linguistics*, *8*(1), 235-43.

Waddington, C. 2001. Different methods of evaluating student translations: The question of validity. *Meta*, *46*(2), 311–25.

Williams, M. 2001. The application of argumentation theory to translation quality assessment. *Meta*, *46*(2), 326-44.

Williams, M. 2004. *Translation quality assessment: An argumentation-centered approach*. University of Ottawa Press, Ottawa.

Williams, M. 2009. Translation quality assessment. *Mutatis Mutandis*. *2*(1). 3-23.

APPENDIX

Table A1: The function-based translation quality assessment.

	Score Range	Description	Rating
1		Accuracy (30%)	
	25-30	There are no identifiable problems of comprehension to the TL readers. The original message is completely transferred to the TT without omissions or additions.	Score: Note:
	21-24	The problems of comprehension are only related to the most highly specialized vocabulary that do not affect the TL readers' understanding. There are some partial omissions and additions.	
	16-20	There is some difficulty experienced by the TL readers in understanding the TT due to the translator's misunderstanding of some parts of original message, apparent omissions and additions.	
	11-15	The ideas are poorly expressed. The translator's serious problems in understanding the ST disturb the transfer of the original message. The TT is difficult to understand.	
	1-10	The translator's serious problems in understanding the ST terribly disturb the transfer of the original message. The TL readers are unable to understand what the original writer intends to convey.	
2		Finding equivalent (25%)	
	20-25	All lexical and syntactic elements are understood through precise vocabulary usage. The words are chosen so skillfully that the translation reads like a good publishable text.	Score: Note:
	15-19	All lexical and syntactic elements are generally understood through good usage of a wide range of vocabulary and structures. Specialized vocabulary presents some problems with inappropriate equivalents.	

	10-14	Most of lexical and syntactic elements are generally understood through a fair range of vocabulary despite some observable gaps. Some vocabulary is inappropriately used. There is some evidence revealing translator's difficulties in finding equivalents, perception, wordplay and other linguistic features.	
	5-9	Comprehension of vocabulary and structures shows quite observable gaps which blur sense. There are problems in finding correct vocabularies due to translator's unability to overcome specialized vocabulary problems.	
	1-4	The vocabulary is inapprpriately used. Translator's poor comprehension of the ST seriously hinders the process of finding equivalents. As a whole, translation makes little sense.	
3	Translation Skill (20%)		
	17-20	The translator demonstrates able and creative solutions to translation problems and is skillful in using resource materials.	Score: Note:
	13-16	The translator demonstrates consistent ability in identifying and overcoming translation problems. The TT contains only very few minor errors. There are no obvious errors in using resource materials.	
	9-12	The translator demonstrates a general ability to identify and overcome translation problems. There is a major translation error and/or an accumulation of minor errors. The improper use of reference materials is possibly reflected in the TT.	
	5-8	The translator demonstrates some difficulty in identifying and/or overcoming translation problems. There are several major translation errors and/or a large number of minor errors. The improper use of reference materials is reflected in the TT.	
	1-4	The TT reflects the translator's inability to identify and	

		overcome common translation problems. There are numerous major and minor translation errors. The TT is seriously flawed translation. The TT contains improper use of materials and resources.	
4	Text/Sentence Function (15%)		
	13-15	Meeting the text/sentence function of the ST; creative inventions and skillful solutions to achieve the function of the ST; corresponding to the text function based on the TL perspective	Score: Note:
	10-12	Almost meeting the text/sentence function of the ST; some inventions to achieve the function of the ST; corresponding to the text function based on the TL perspective	
	7-9	Inconsistency in meeting the text/sentence function of the ST; awkward structure in achieving the ST function; not fully corresponding to the text function based on the TL perspective	
	4-6	Less attention to the text/sentence function of the ST; not corresponding to the text function based on the TL perspective	
	1-3	In contrast to the text/sentence function of the ST; not corresponding to the text function based on the TL perspective	
5	Grammar and TT style (10%)		
	9-10	The TT needs no improvement in terms of its grammatical and stylistic points though one or two natural failings are possibly observable. The TT shows native-like fluency in grammar.	Score: Note:
	7-8	The TT needs a little improvement in terms of its grammatical and stylistic points. The TT maintains advanced proficiency in grammar; There are some grammatical problems, but they do not affect the message.	
	5-6	The TT tends to have awkward grammatical usage due to literal rendering of the ST message, but it does not	

		significantly hinders the message transfer. There are some attempts to reflect stylistic features of the TL. Some grammatical problems are apparent and have negative effects on message transfer.	
	3-4	The TT is awkward due to nonsensical grammatical usages. The TT sounds unnatural. There is a little attempt to reflect stylistic features of the TL. There is evidence of clear difficulties in following the TL style. Grammatical review of some areas is obviously required.	
	1-2	Little sense of the TL style. Knowledge of TL grammar is inadequate. The use of TL grammar is inadequate. Serious grammatical problems truly affects the message transfer.	