

## Relationship between Two Translation Quality Assessments: Holistic Rating and Waddington's Model of Assessment



Maryam Asgari<sup>1\*</sup>

<sup>1</sup>English Department, Islamic Azad University, Bandar Abbas, Iran

### Citation

Asgari, M. (2021). Relationship between Two Translation Quality Assessments: Holistic Rating and Waddington's Model of Assessment. *International Journal of Language and Translation Research*, 1(2), pp. 27-46.  
DOI: 10.12906/978389966720\_002

### Abstract

#### Available online

#### Keywords:

Holistic method,  
objective  
translation,  
assessment criteria,  
translator training,  
Waddington's  
Model

Translation quality assessment is one of the most significant and, at the same time, problematic areas of translation. The critical importance of this issue becomes more obvious in pedagogical contexts. The present study focused on the translation quality assessment undertaken in Islamic Azad University of Bandar-Abbas which offers translation training in both B.A and M.A levels. In this study, Waddington's model of TQA, which is accepted as an objective model, was applied to the exam papers of the students, already assessed and scored by their instructors. The results obtained from statistical analysis of the data, that is, the two sets of scores, revealed that a correlation does exist between the scores obtained through applying Waddington's model and the scores assigned to the papers by the instructors. Based on this finding, two conclusions were drawn: 1) the assessment carried out in the above-mentioned university is objective, and 2) Waddington's model and its criteria are not that much objective, and has some shortcomings. One of the shortcomings, according to the findings of the present study, is that 'the unit of translation' has not been specified in the model. Thus, the researcher proposes to consider concept' as the 'unit of translation'.

<sup>1</sup>Corresponding Author's Email:  
[m.asgari2139@gmail.com](mailto:m.asgari2139@gmail.com)

### Introduction

Translation is a hermeneutic process in which intuition plays a crucial role in interpreting the intentions of the source text writer (Al-Qinai, 2000 c.f. Khanmohammad, 2009). Translation is a work of art and should be treated as an artistic field of study. So, the term translator cannot be attributed to all the people transferring concepts from one language into another. The translated text is a piece of art. The masterpiece created by an artist named translator. As Pym (2009:1) states: “Just as everyone can sing, be it badly or well, so everyone who knows more than one language can translate, to some degree. However, not everyone is paid to sing opera, and not all translators are at the principal of the translation profession.”

But how are translators different from each other? According to Pym (2009), the difference between various levels may partly be due to training (P.1). What is the aim of training? What do we expect from a translator we are training? He adds, “We train people not just to translate, which they can already do, but to translate well” (Pym, 2009: 2). Heydari Tabrizi (2008:1) believes that translator training is gaining increase in the world and also in Iran. He says: “During the last decade or so, the number of Iranian universities offering the academic Translation Program ... has been increasing.”

The purpose of the present research is to find out whether the assessments undertaken in Iranian universities are objective or not. The present study has been conducted on Islamic Azad university of Bandar-Abbas. The reason for selecting this university was that translation is being taught there both in B.A and M.A levels. Therefore, it is one of the universities specialized in this field. The students graduated from this university are supposed to have acquired the knowledge of professional translation and the application of the technical procedures during the task of translation. To do so, all the aspects of training should be taken into consideration: the subject matter to be taught, the instructional schedule, adequate number of practical courses, and the last but not the least, training assessment. According to Mokolič Južnič (2013), “Assessment is an essential part of training”. Of course, so many investigations have been conducted in all the above-mentioned aspects in this university, but it seems that some more investigations are needed in the assessment area. Mokolič Junžnič (2013) assumes that assessment as one of the trainers’ jobs is not only for obtaining grades in a single course, but also is a way to help learners improve their skills. “The feedback it provides is essential for the development of their skills” (Mokolič Južnič,

2013). Here a question arises: if the feedback of the instructors' assessment is important as such, what happens if the assessment would not be objective enough?

### **Background of the Study**

According to Pym (2009), the "traditional" translation class has been entirely unprofessional as the trainers were asked to apply the traditional "didactic translation" (p. 5) model in which the students were to provide texts only for the teacher to read, in a way that the translation is evaluated positively only when it corresponds to the way the teacher translates" or just like the teacher's translations in the class.

There used to be different schools of thought in translation evaluation. The "Mentalist views" which regard no core meaning for the texts and believe that their meaning alter depending on individual speakers' positions. The "Response Based Approaches which are divided into two categories; namely, behavioristic views and functionalistic, skopos-related approaches. The former "takes reader reactions to the translation as the main yardstick for assessing a translation's quality" (House, 2001:244) and in the latter, the "purpose" of the translation is of the great importance.

The third main school of thought is "Text and Discourse-Based" approaches which has three categories named literature-oriented approaches, post-modernist and deconstructionist thinking and linguistically oriented approaches.

Subsequent to these approaches, a new season began in translation evaluation which presented a grading scale for the trainers. These models which use error analysis, assign a specific value for each kind of error distinguished in the translations. These kinds of assessment models are more appropriate for pedagogical contexts, as one of the aims for the assessment in academic context is obtaining scores for translation courses.

So many investigations and attempts have been made to improve the quality of translation programs especially "the often-neglected" part of it which is "teacher evaluation of the trainee translations" (Haydari Tabrizi, 2008). Haydari Tabrizi assumes the subjective nature of translation evaluation as the reason for its difficulty and tricky status. It means that the quality is a notion which changes from one situation to another and one evaluator's point of view from another evaluators. But such an evaluation cannot help the trainees. They cannot make any decision about

their own weaknesses and improvements, because some errors on one evaluator's point of view may not be considered as errors for another one.

Malcom Williams (2001:328 c.f. Sabiza, 2009) believes that "there are no generally accepted objective criteria for evaluating the quality of both translation and interpreting performance."

There are many studies which have concluded that the assessments undertaken in Iranian universities are not objective. Sabiza (2009:15) believes that teachers in Iranian universities apply subjective criteria for assessing and grading the trainees' translations. She supposes that grading scales are moderately used and "each trainer uses his or her own criteria for evaluation". She supposes that the trainer in the pedagogical context is a judge, the trainee has to submit his or her authority which may be just and objective or not.

Most of Iranian instructors participated in Khanmohammad's (2009) study, have advocated "the possibility of objective assessment of the students' translations, but she attributes this to the fact that "they suggest typical equivalents and translations for words and texts in the class and they expect students to emulate the same translation at the time of exam" (Khanmohammad, 2009:143). She supposes that this might be the notion that makes the Iranian instructors believe the students' translation can be assessed objectively.

This issue may result in many problems in the process of training. First of all, the students cannot trust their scores. It can be higher or lower if someone else evaluates their translation. Besides, the scores obtained are not representative of the students' knowledge and skill. Therefore, if the above-mentioned arguments would be right in all the Iranian universities, it can be said that it is not a valid measurement of the students' learning.

The present study examined such argumentations. The researcher chose Islamic Azad University of Bandar abbas for the above-mentioned reasons to investigate the assessment procedure applying in it and observe if a valid and objective assessment is done there or not. In order to achieve this goal, an objective model of translation quality assessment was selected to apply to the already assessed translations and observe if the scores obtained through this model correlates with the scores given to the students' translations by their own instructors.

So many investigations have been conducted in the field of assessment and most of them have applied the existing objective models. So many criticisms have also been undertaken on literary

works by objective models, (for example, Shahraki and Karimnia, 2011), but the number of studies on the assessment of translations in academic contexts are just a few.

The significance of this study lies in the fact that translation trainees are the potential translators who will arrive into translation industry as professional translators in near future. Therefore, their training program should be professional enough. This study has actually worked on the last but not the least layer of translation instructional program which is assessment. The assessment procedure too was undertaken in one of the universities in Iran in which Translation Studies and specialized translation courses are offered at both B.A and M.A levels.

The purpose of this study, as was mentioned earlier, was to see whether the assessments applied in Islamic Azad University are objective or subjective. Actually, the study was to find out if the measurements undertaken by the instructors are representative of the translation students' trainees' knowledge. Thus, the following question was addressed and the related null hypothesis was tested:

### **Research Question and Hypothesis**

**RQ.** Is there any correlation between the translation instructors' scores and the obtained scores by reassessing the exam papers through Waddington's model of translation quality assessment?

**HO.** *There is not any correlation between the translation instructors' scores and the scores obtained by reassessing the exam papers applying Waddington's model of translation quality assessment.*

### **Method**

#### **Participants**

For the purposes of this study, the researcher collected 200 examination papers of the course "Advanced translation" belonging to five classes between the years 2013 to 2015, belonging to male and female students who had been in their 8th semester of study at university level. Out of the total papers, 60 papers were randomly selected to be reassessed on the basis of Waddington's model.

#### **Materials**

The materials of the study consisted of the texts of the tests taken and scored. Each test used in this study consisted of two texts. One was an English text to be translated into Persian and the

other was a Persian text which was given to the participants to be translated into English. The mean length of the English texts was 117 words and the mean length of the Persian texts was 75 words. All the texts included two or three paragraphs. Therefore, the texts were approximately of the same length.

### **Level of Difficulty of the Texts**

The difficulty of the texts was measured through the system used by Waddington (2003) as follows:

- i) Degree of necessary re-expression (R)
- ii) Number of translation problems (Pii)
- iii) Number of lexical differences (LD)
- iv) Number of syntactic differences (SD)
- v) Failure to fulfil the learning objectives (LO)

Based on the above measurement, the degree of difficulty for the texts was almost the same ( $SD=1.4$ ).

### **Model of the Study**

The following model was used in this study to reassess the exam papers of the students.

### **Waddington's Model**

This model consists of the following 4 methods for the assessment of translation papers:

#### **Method A**

Method A is taken from Hurtado (1995); it is based on error analysis, and possible mistakes are grouped under the following headings:

(i) Inappropriate renderings which affect the understanding of the source text; these are divided into eight categories: *contresens*, *faux sens*, *nonsens*, addition, omission, unresolved extra linguistic references, loss of meaning, and inappropriate linguistic variation (register, style, dialect, etc.).

(ii) Inappropriate renderings which affect expression in the target language; these are divided into five categories: spelling, grammar, lexical items, text and style.

(iii) Inadequate renderings which affect the transmission of either the main function or secondary functions of the source text.

In each of the categories a distinction is made between major errors (–2 points) and minor errors (–1 point). There is a fourth category which describes the plus points to be awarded for good (+1 point) or exceptionally good solutions (+2 points) to translation problems. In the case of the translation exam where this method was used, the sum of the negative points was subtracted from a total of and then divided by 11 to reach a mark from 0 to 10 (which is the normal Spanish system). For example, if a student gets a total of –66 points, his result would be calculated as follows:  $110 - 66 = 44 / 11 = 4$  (which fails to pass; the lowest pass mark is 5).

### Method B

Method B is also based on error analysis and was designed to take into account the negative effect of errors on the overall quality of the translations (Kussmaul 1995:129). The corrector first has to determine whether each mistake is a translation mistake or just a language mistake; this is done by deciding whether or not the mistake affects the transfer of meaning from the source to the target text: if it does not, it is a language error (and is penalized with 1 point); if it does, it is a translation error (and is penalized with –2 points). However, in the case of translation errors, the corrector has to judge the importance of the negative effect that each one of these errors has on the translation, taking into consideration the objective and the target reader specified in the instructions to the candidates in the exam paper. In order to judge this importance, the corrector is given the following table:

**Table 1**

*Typology of Errors in Method B*

<i>Negative effect on words in St</i>	<i>Penalty for negative effect</i>
On: 1-5 words	2
6-20 words	3
21-40 words	4
41-60 words	5
61-80 words	6
81-100 words	7

100+ words	8
The whole text	12

The final mark for each translation is calculated in the same way as for Method A, that is to say, the examiner fixes a total number of positive points (in the case of method B, this was 85), then subtracts the total number of negative points from this figure, and finally divides the result by 8.5. For example, if a student is given 30 minus points, his total mark would be 6.5 (pass):  $85 - 30 = 55 / 8.5 = 6.5$ .

### **Method C**

Method C is a holistic method of assessment. Although, in the survey mentioned above, the teachers who answered were requested to send a brief description of the method of assessment they applied, I only received three descriptions of holistic methods. In addition to this, all three methods based their scales on the requirements of professional translation and were consequently of little use for judging the quality of translation into the foreign language. As a result, I had to design the following holistic method myself. The scale is unitary and treats the translation competence as a whole, but requires the corrector to consider three different aspects of the student's performance, as shown in the table below. For each of the five levels there are two possible marks, so as to comply with the Spanish marking system of 0 – 10; this allows the corrector freedom to award the higher mark to the candidate who fully meets the requirements of a particular level and the lower mark to the candidate who falls between two levels but is closer to the upper one.

### **Method D**

Method D consists of combining error analysis Method B and holistic Method C in a proportion of 70/30; that is to say, Method B accounts for 70% of the total result and Method C for the remaining 30%.

### **Why Method A?**

In the present study, the researcher preferred Method A to Method B because of three reasons. The first one is that Method B considers the effect of the each error on the overall quality of the



translation and this effect is determined based on the objective of translation and the target reader (Waddington, 2003), but the translated texts in translation classes are just for evaluating the students' skill and their knowledge of translation, and the translated texts are not going to be published; so there is no target reader.

**Table 2**

*Scale for holistic Method C*

Level	Accuracy of transfer of ST content	Quality of expression in TL	Degree of task completion	Mark
Level 5	Complete transfer of ST information; only minor revision needed to reach professional standard.	Almost all the translation reads like a piece originally written in English. There may be minor lexical, grammatical or spelling errors.	Successful	9, 10
Level 4	Almost complete transfer; there may be one or two insignificant inaccuracies; requires certain amount of revision to reach professional standard.	Large sections read like a piece originally written in English. There are a number of lexical, grammatical or spelling errors.	Almost completely successful	7, 8
Level 3	Transfer of the general idea(s) but with a number of lapses in accuracy; needs considerable revision to reach professional standard.	Certain parts read like a piece 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	Transfer undermined by serious inaccuracies;	Almost the entire text reads like a translation; there are	Inadequate	3, 4

The second reason can be found from a study conducted by Waddington (2003) that shows Method B is less applicable for translating the text into a foreign language. As the researcher aimed at checking both English to Persian and Persian to English translation, she had to choose a method which was appropriate for both types of translation. And the third problem of Method B was the

time it consumes. In an investigation conducted by Tamara Mikolič Južinič, it was stated that 31.5% of the trainers find their “assessment system functional but time consuming”. It can be inferred from this statement that the trainers prefer methods which are functional but take less time (Mokolič Južinič, 2013).

Waddington (2003), comparing these two methods with Methods C and D, concludes that “the two error analysis methods produce more consistent results than the holistic method”. He logically supposes dangerous to reduce the overall quality of a student translation by summing up the mistakes encountered. Therefore, Methods C and D were rejected too and Method A was chosen. Thus, the researcher prepared a checklist based on Waddington’s TQA model, Method A, which is taken from Hurtado (1995).

### **Procedure**

As mentioned above, a total of 200 exam papers of the students belonging to the course of Advanced Translation comprised the material of the present study, out of which 60 papers were selected randomly for careful scrutiny. Then, the scores given to these papers by instructors and the scores obtained from them through reassessment on the basis of Waddington’s model were compared and statistically analyzed for possible relationship between these two modes of scoring. To this purpose, the researcher prepared and used a checklist based on the instructions of the model of the study.

It should be mentioned that effort was made to select the exam papers for reassessment and rescoring that had texts with almost the same level of difficulty ( $SD= 1.4$ ). The difficulty level was measured on the basis of Waddington’s (2003) criteria as follows:

- i) Degree of necessary re-expression (R)
- ii) Number of translation problems (Pii)
- iii) Number of lexical differences (LD)
- iv) Number of syntactic differences (SD)
- v) Failure to fulfil the learning objectives (LO)

Finally, to make sure of the validity of the scores obtained by the researcher, the selected exam papers were assessed by three raters, all of whom Ph.D. graduates of English translation. They

were asked to assess the papers based on Waddington's model, Method A. As for reliability, each paper was assessed three times by the raters.

### Results

In this section, the statistical analysis of the data and the obtained results are presented. As mentioned before, the students' papers, previously rated by translator instructors, were assessed and scored by three raters using Waddington's Model. Each paper was rated three times by a rater and the mean of the three ratings for each paper was taken as the final score given to that paper. This was done to make the ratings more reliable and decrease the effect of possible interfering variables. Additionally, the researcher herself, rated each paper once, using Waddington's Model. Table 3 below presents the related descriptive statistics:

**Table 3**

*Descriptive Statistics for All Scores (N=60)*

	<i>M</i>	<i>SD</i>
Instructors' Scores	6.29	2.27
Model Scores	7.49	1.60
Overall Rater Scores	7.17	1.60
First Rater's Scores	7.18	1.61
Second Rater's Scores	6.97	1.51
Third Rater's Scores	7.37	1.79

As seen in the Table above, while mean of the overall rater scores ( $M=7.17$ ) is higher than the mean of instructors' scores ( $M=6.29$ ), the standard deviation of instructors' scores ( $SD=2.27$ ) is higher than that of the overall rater scores ( $SD=1.60$ ). In addition, mean of the model scores is 7.49 and their standard deviation is 1.60.

### Reliability of Rating Scores

To make sure that raters' ratings are reliable, it was necessary to check for both inter-rater and intra-rater reliability of scores given by raters.

### Intra-rater Reliability

This measure of reliability was checked for each rater separately. Table 4 below presents correlations between three sets of scores given to papers by the first rater.

**Table 4**

*Correlations between the First Rater's Ratings*

	First	Second Rating	Third Rating
First Rating	-	.96	.95
Second Rating	.96	-	.98
Third Rating	.95	.98	-

Table 5 below shows the correlations between the second rater's ratings:

**Table 5**

*Correlations between the Second Rater's Ratings*

	First Rating	Second Rating	Third Rating
First Rating	-	.89	.86
Second Rating	.89	-	.97
Third Rating	.86	.97	-

According to Table 3, the highest correlation is between the second and the third ratings ( $r=.97$ ) and the lowest is between the first and the third ratings ( $r=.86$ ). Using Spearman-Brown prophecy formula, intra-rater reliability for the second rater turned out to be .94. Table 6 depicts the correlations between ratings of the third rater.

**Table 6**

*Correlations between the Third Rater's Ratings*

	First Rating	Second Rating	Third Rating
First Rating	-	.99	.95
Second Rating	.99	-	.95
Third Rating	.95	.95	-

Here in Table 4, the highest correlation is between the second and the first ratings ( $r=.99$ ) and correlations between the first rating and the third and the second rating and the third are both  $.95$ . Calculation of intra-rater reliability using Spearman-Brown prophecy formula for this rater gives us a result of  $.98$  that is equal to the intra-rater reliability index of the first rater.

### Inter-rater Reliability

Having made sure that intra-rater reliability is high enough, it's necessary to check for inter-rater reliability as well. The procedure for obtaining inter-rater reliability is straightforward and similar to the procedure for obtaining intra-rater reliability.

First, correlations between overall scores given by each rater have to be obtained. Table 7 presents these correlations.

**Table 7**

*Correlations between Overall Rating Scores of Raters*

	First Rating	Second Rating	Third Rating
First Rating	-	.95	.94
Second Rating	.95	-	.89
Third Rating	.94	.89	-

Using Spearman-Brown prophecy formula, inter-rater reliability turned out to be  $.96$  which indicates the very high inter-rater reliability of the ratings.

### Testing Research Hypothesis

The hypothesis of this study states that,

*There is not any correlation between the translator instructors' scores given to the students with the scores obtained by reassessing the exam papers applying Waddington's model of translation quality assessment.*

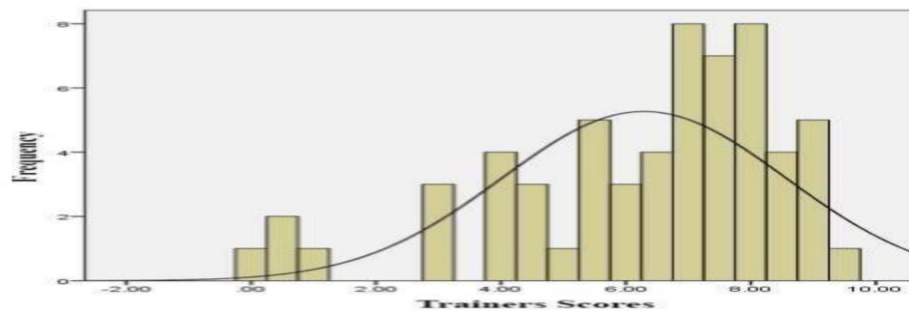
To test this hypothesis, Pearson's correlation is the statistical test of choice, however, it's necessary to check for possible violations of the assumptions underlying Pearson's correlation.

### Assumption Testing

Three main assumptions of Pearson's correlation were checked before statistical testing: normality, linearity, and outliers. To check for normality, histograms for three main variables in this study (i.e. Instructors' Scores, Model Scores and Rating Scores) were checked. Figures 1, 2, and 3 below depict histograms for instructors' scores, model scores and rating scores respectively.

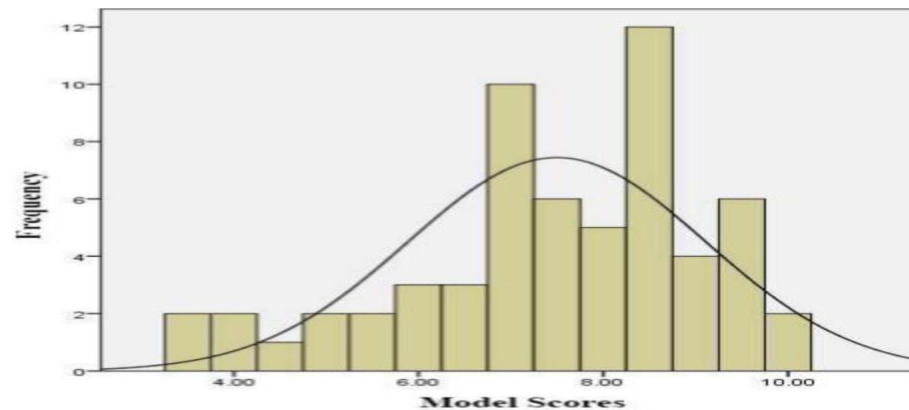
**Figure 1**

*Histogram for Instructors' Scores (N=60, M=6.29, SD=2.27, Skewness=-1.05)*



**Figure 2**

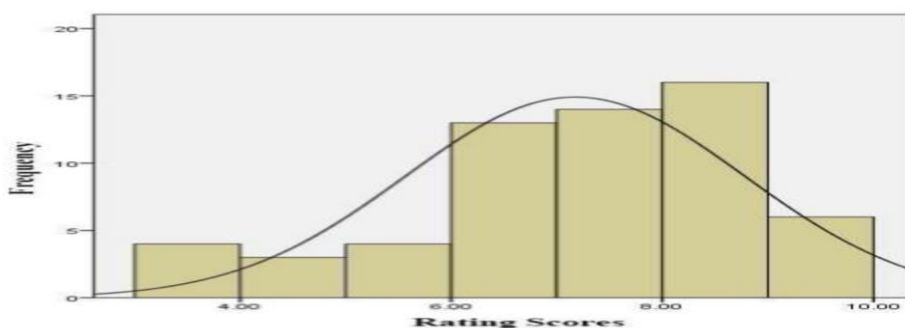
*Histogram for Model Scores (N=60, M=7.49, SD=1.60, Skewness=-.82)*



All figures above show that scores are negatively skewed. Skewness values of minus 1.05, minus .82 and minus .80 show that, to some extent, the assumption of normality has been violated. In order to look for possible violations of the assumption of linearity and to locate outliers, scatterplots for the interaction of rating scores and model scores and the interaction of instructors' scores and model scores needed to be checked. Figures 4 and 5 show the scatterplots for the above-mentioned interactions.

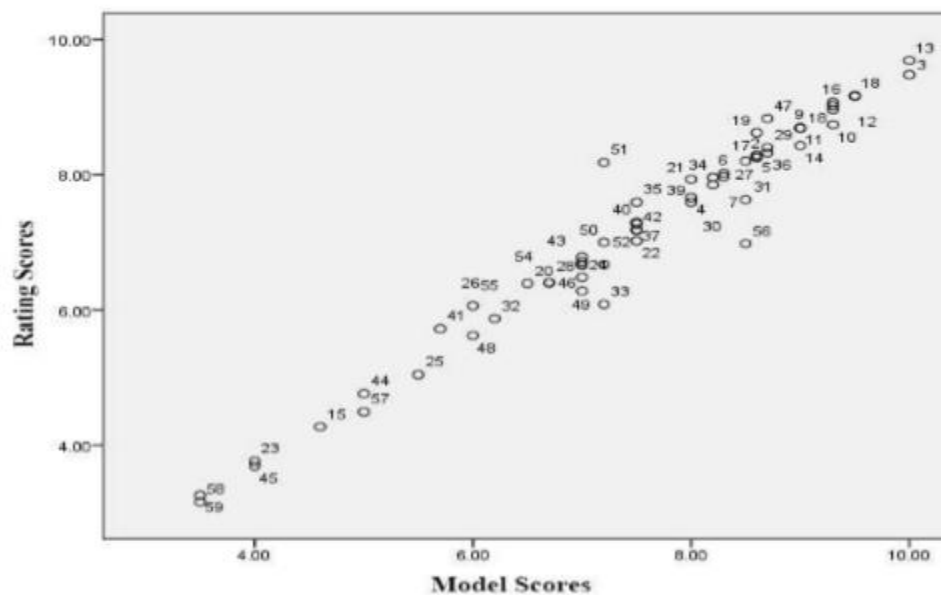
**Figure 3**

*Histogram for Rating Scores (N=60, M=7.17, SD=1.60, Skewness=-.80)*



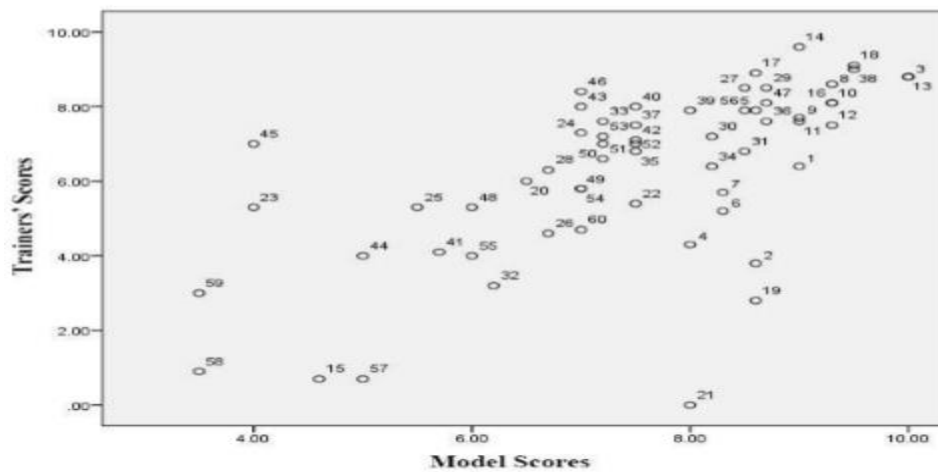
**Figure 4**

*Scatterplot for the Interaction of Rating Scores and Model Scores*



**Figure 5**

*Scatterplot for the Interaction of Instructors' Scores and Model Scores*



Looking at the scatterplots in Figures 4 and 5, no evidence of a curvilinear relationship can be found; hence, it can be argued that the assumption of linearity has not been violated. Regarding outliers, while some outliers can be seen in the scatterplots, especially in Fig 5, as the number of outliers is small, it is unlikely that they affect the results, and it was decided to retain them.

### **Correlation between Overall Rater Scores and Model Scores**

Before testing our research hypothesis, it is necessary to check for the correlation between overall scores obtained by assessing papers using Waddington's model by three raters and the scores obtained by assessing papers using Waddington's model by the researcher herself. However, in this case, Pearson's correlation is the statistical test of choice, due to aforementioned violations of the assumption of normality, it's necessary to use Spearman's correlation as the non-parametric equivalent of Pearson's correlation. Table 6 below presents the results of correlation analysis.

**Table 6**

*Correlations between Model Scores and Overall Rating Scores*

Model Scores	Overall Rating Scores	
	<i>rho</i>	<i>P</i>
	.97	.00



According to Table 6, these two sets of scores are highly correlated ( $\rho=.97$ ,  $p=.00$ ). This means that the results of ratings using Waddington's model by the three raters is quite close to the result of ratings using Waddington's model by the researcher herself.

### Correlation between Model Scores and Instructors' Scores

As mentioned earlier, due to violations of the assumption of normality, it was decided to use Spearman's correlation as the non-parametric equivalent of Pearson's correlation. Table 7 below shows the results of Spearman's correlation for the interface of model scores with translator instructors' scores, conducted to test our research hypothesis concerning the possible interface of scores given by translator instructors and scores given by the researcher herself using Waddington's model.

**Table 7**

*Correlations between Model Scores and Instructors' Scores*

Model Scores	Translation Instructors' scores	
	<i>rho</i>	<i>P</i>
	.64	.00

According to Table 7, the correlation coefficient for the relationship between model scores and instructors' scores is .64. *P* values for this correlation is .00, which is lower than our alpha level ( $\alpha=.05$ ); hence, it can be said that the null hypothesis of the study is proven to be wrong. In other words, translation instructors' scores are significantly correlated with the scores obtained through reassessing the exam papers applying Waddington's model of translation quality assessment. The following section deals with whys and wherefores of this relationship.

### Discussion

This study sought to find out if the translation quality assessment undertaken in Islamic Azad University of Bandar Abbas is objective or not. In order to do that, the already assessed translation exam papers were reassessed through Waddington's model of TQA, which is known as an objective model. Then, correlation was calculated between the two sets of scores- the scores given to the students' translations by their instructors on the one hand and the scores obtained through

applying Waddington's model of TQA, on the other. To make sure of the validity of the scores obtained, the exam papers were assessed three times by three rates, all of whom were Ph.D. graduates of English Translation.

Although Pearson correlation was the statistical test of choice for testing the hypothesis, when the three assumptions of Pearson correlation; namely, normality, linearity and outliers were checked, it was found that the assumption of normality had been violated, but the two others had been saved. Therefore, Spearman correlation was used instead. Calculating the Spearman's correlation coefficient, it was found that there is a significant correlation between the instructors' scores and those obtained through applying Waddington's model of TQA ( $\rho=64$ ,  $p<0.05$ ). This means that the hypothesis assuming that there is no relationship between the instructors' scores and those obtained through applying Waddington's model is wrong.

### Conclusions

As it was seen, the correlation did exist between the two sets of scores. Thus, based on the obtained results, two conclusions can be inferred. First, the instructors' scores correlate with an objective model of TQA, and so the instructors' evaluation is objective too and opposed to the claims of such researchers as Haydari Tabrizi (2008) and Sabiza (2009). Haydari Tabrizi (2008:3) argues that the validity, reliability, practicality and even the way of grading of Iranian instructors is under serious question. He asserts that "translation teachers of Iranian universities are least informed and familiar, if at all, with the current translation evaluation approaches". He adds that the dominant trend for translation quality assessment in Iran is far behind the modern ones practiced in accredited universities throughout the world" (p.3). Sabiza (2009:2) too assumes that the teachers apply subjective criteria for their assessment and grading of students' translations. These researchers' arguments might be right according to the population they have worked on at that time of conducting research, but cannot be generalized to all the Iranian universities, as the present study rejects them.

This investigation showed that the instructors in Islamic Azad University of Bandar Abbas have followed the modern strategies of translation quality assessment and the objective models presented during the last decade. Therefore, their evaluation can be a measure for the students'

competence, and therefore, this university has fulfilled one of the requirements of a professional instruction in the field of translation; that is, *objective assessment*.

The second conclusion inferred from the obtained results, could be just on the opposite. Observing correlation between the scores based on Waddington's model of TQA and the scores given to the students by their instructors may mean that the Model is not that much objective. It means that although the most objective method proposed in the Model; namely, Method A, which is based on error analysis, was selected for this study, it was found that the unit of translation had not been exactly determined in the method. For example, in case of "omission", Waddington (2001) has assigned -1 points penalty for minor omissions and -2 points penalty for serious omissions, while the unit of omission has not been defined. There are some objective models like Farahzad's (1992) model of TQA which maintains that "sentence and clause might be the unit of translation" (Farahzad, 1992 c.f. Khanmohammad, 2009:4). She elaborates that "each verb in the source language text marks a score" (Farahzad, 1992 c.f. Khanmohammad, 2009:4). But there is not such explication in Waddington's model. If a paragraph or even a sentence has not been translated at all, how many points should be subtracted based on Waddington's model? Of course, the researcher does not believe that Farahzad's assertion is completely appropriate; as we have smaller units in a sentence or clause which should be taken into consideration. Instead, the researcher introduces 'concept' as the unit of translation, since every concept can be carried by a lexical item, a grammatical indicator, e.g. plural "s" the third person indicator, "s" or "es", etc. For example, in the sentence 'the rain tapped against the window', the concept of 'lightly' is the implicit meaning of the verb 'tap', and it should be transferred during the process of translation whether implicitly or explicitly depending on the capacities of the target language.

### References

- Al-Qinai, J. (2000). Translation Quality Assessment. Strategies, Parameters and Procedures. *Translators' Journal*, 45(3), 497–519.
- Farahzad, F. (1992). "Testing achievement in translation classes", in C. Dollerup and A. Loddegaard (eds.), *Teaching Translation and Interpreting: Training, Talent, and Experience*, Amsterdam/Philadelphia: John Benjamins Publishing Company, pp. 271-278.

- Heydari Tabrizi, H. (2008). Towards Developing a Framework for the Evaluation of Iranian Undergraduate Students' Academic Translation. M.A. Thesis. University of Shiraz. Available from [www.irandoc.ir](http://www.irandoc.ir) [Accessed 18 Jan 2016].
- House, J. (2001). Translation Quality Assessment: Linguistic Description versus Social Evaluation. *Meta: Translator's Journal*. 46(2).
- Khanmohammad, H, and Osalnoo, M. (2009). Moving Toward Objective Scoring: A Rubric for Translation Assessment. *JELS*. 1(1). 131-153.
- Kussmaul, Paul, Sonja Tirkkonen-Condit (1995). Think-aloud protocol analysis in translation studies. *TTR: traduction, terminologie, rédaction* 8: 1, 177–199. Available at <http://id.erudit.org/iderudit/037201ar> accessed on August 21, 2011.
- Mokolič Južnič, T. (2013). Assessment Feedback in Translator Training: A Dual Perspective. *Publications of the University of Eastern Finland Reports and Studies in Education Humanities and Theology*. 8(3). 74-101.
- Pym, A. (2009). *Exploring Translation Theories* (London and New York: Routledge).
- Sabiza, S. (2009). The Iranian University Teachers' Criteria for Evaluation of Students' Translation. M.A. Thesis. Available from [www.irandoc.ir](http://www.irandoc.ir). [Accessed 19 Jan 2016].
- Shahraki, A., & Karimnia, A. (2011). Waddington's Model of Translation Quality Assessment: A Critical Inquiry. *Elixir Ling. & Trans.* 40. 5219-5224.
- Waddington, C. (2003). A Positive Approach to the Assessment of Translation Errors. *AIETI*. 2. 409-426.