INTERNATIONAL JOURNAL OF

LANGUAGE AND TRANSLATION RESEARCH

1 (2021) 2

International Journal of Language and Translation Research (IJLTR) is a peer-reviewed, quarterly print/online journal with an editorial board of scholars in the fields of language teaching, linguistics, literature, and translation studies from different parts of the world. It welcomes the submission of research-based articles and reviews on various aspects of English language teaching/learning and translation. Submissions should comprise relevant theoretical foundations and pedagogical implications. They should further considerably contribute to related literature existing.

Ethical Statement

The Ethical Statement is based on the recommendations of the Publication Ethics Committee (COPE) Good Practices drafted in 2011.

1. Obligations of the editor:
   1.1. Neutrality. The intellectual content of submitted manuscripts is evaluated is evaluated regardless of race, gender, sexual orientation, age, disability, religion, ethnicity, political philosophy of the authors.
   1.2. Confidentiality. All manuscripts should be treated as confidential documents. They must not be shown to anyone without the permission of the editor. Managers and editorial staff should not disclose information about the manuscript submitted to anyone except the author, reviewers and potential reviewers.
   1.3. Disclosure of information and conflicts of interest. Unpublished data contained in the submitted manuscript must not be used by editors or reviewers in their own research without the explicit consent of the author.
   1.4. Decision on publication. The editor of the journal decides on the publication of submitted articles. The editor is guided by the Editorial Committee's policy, taking into account the legal obligations regarding defamation, copyrights and plagiarism. The editor can share the decision with other members of the Editorial Board or with reviewers. In the event of an appeal of the decision of the Reading Committee, the editor may solicit two new reviewers.
2. Obligations of reviewers.
   2.1. Editorial decisions. Reviewers assist the editorial staff in making decisions and may also assist the author to improve the quality of the manuscript.
   2.2. Delays and deadlines. When a guest reviewer does not feel competent enough to evaluate the research presented in the manuscript, or if he finds himself unable to provide his report in time, he must inform the editor without delay in order to give him time to contact other reviewers.
   2.3. Standards of objectivity, civility and respect. The reports must be objective. Personal remarks and criticisms directed at the author or hurtful remarks directed at the text content are not eligible. The opinion of the reviewer must be clear, well-argued and respectful of the author.
   2.4. Indication of sources. The reviewer must identify appropriate publications not cited by the author. Any such indication must be accompanied by an appropriate comment. The reviewer should draw the editor's attention to any similarity, any overlap between the manuscript and previously published data.
   2.5. Disclosure of information and conflicts of interest. Information and ideas obtained through anonymous replay are confidential and should not be used for the personal benefit of the reviewer. Reviewers should not accept reviewing manuscripts where this may result in a conflict of interest arising from competitive, collaborative or other relationships with the authors.
3. Obligations of the authors.
   3.1. Information validity. The information contained in the manuscripts submitted for publication must present the results of the authors' research as well as an objective discussion of these results and their importance. The underlying data must be presented correctly. Fraudulent and consciously inaccurate information is considered unethical and unacceptable. The identification of research done by others must always be given. Authors should cite the publications that influenced the study in question.
   3.2. Originality and plagiarism. Authors must ensure that they have written a completely original study, and if they have used other people's books or statements, they must be properly cited.
   3.3. Multiple publications. An author should not submit manuscripts representing the same study to more than one journal (or book). Submitting the same manuscript in more than one journal is unethical and unacceptable. The journal accepts articles originally published in languages other than English. In these cases, the authors must give the reference of the first publication and be free from the copyright of the original publisher.
   3.4. Paternity of the manuscript. Only authors who have made a significant contribution to the study in question are considered to be authors. All those who contributed to the study must be present in the list of authors. If other people have been involved in some aspects of the research project, they should be mentioned in the acknowledgments. The lead author must ensure that all co-authors and only they are included in the list of authors of the manuscript, that the co-authors have seen and approved the final version of the manuscript, and that they have agreed to submission of the manuscript.
   3.5. Disclosure of information and conflicts of interest. All authors must indicate, as a result of their biographical presentation, any conflicts of interest that may affect their proposed publication. Funding for research projects that made the study possible must be indicated.
   3.6. Errors in publishing. If the author discovers an important error or an inaccuracy in its publication, its obligation is to quickly inform the editor and to consider, in agreement with the person in charge, the withdrawal of the article or the publication of the information about the error.

Effect of Textual Integrity of Argumentative Texts on EFL Learners' ReadingPerformance: Different Levels of Language Proficiency in Focus

Mona Lavenezhad[1*]

[1]English Department, Najafabad Branch, Islamic Azad University, Najafabad, Iran

Abstract

The present study aimed at investigating the effect of textual integrity of argumentative texts on EFLlearners' reading comprehension performance. It also aimed at checking the extent of such an effect among learners with different language proficiency. To this purpose, 120 students learning English at Jihad Daneshgahi Institute in Isfahan were selected as the participants. They were selected from apool of 200 available and interested students and were divided into three groups of low proficient, intermediate and high proficient learners of equal size (40), based on their scores on an OQPT proficiency test. Then, 3 reading comprehension tests (cloze tests), with an appropriate level of text difficulty, were prepared by the researcher. In making the tests, the text in each test was either kept authentic in terms of textual integrity (i.e. text organization, cohesive devices, etc.), or manipulated to lose its textual unity and, thus, be more difficult to read and understand. The results of data analysisindicated that manipulated argumentative texts negatively affect EFL learners' reading performanceat all levels of language proficiency. The results additionally revealed that text manipulation, i.e. textual integrity decrease, has a more significant effect on the reading performance of the intermediate group participants. The findings of this study can have some implications for language teachers to become more alert to the effect of textual integrity of texts on reading comprehension performance of students when trying to understand argumentative texts. Furthermore, the findings might be constructive for materials developers, i.e helping them to prepare appropriate texts in terms of textual integrity and readability, in line with the needs and levels of EFL learners.

[1] Corresponding Author's Email:
mana.lavin@gmail.com

Introduction

Reading comprehension is one of the main objectives of teaching English in an EFL context and it isthe most tested construct in language teaching. The importance of reading comprehension is underscored in today's "information age" in which the ability to read easily and well has become a survival skill: reading "has been considered one of the skills required of all language learners" (Chastain, 1988, p. 2). It is both a source of information and a pleasurable activity, the one which serves as a vehicle for communication of present and past civilizations, and which many students have an opportunity to use (Rivers, 1968; Chastain, 1971). Chastain states that "one of the basic andcomplementary skills which need to be acquired in foreign language learning is reading" (p. 6). Anderson (2001) even goes beyond this and claims that "reading is all that is needed by learners of English as a foreign language (EFL)". In fact, there is a direct relationship between learning a new language and reading, as Bugel and Bunk (1990) assert: "where there is little reading, there will be little learning." (p.17).

Extensive research has been conducted to investigate the relationship between reading comprehension and other areas like semantics, pragmatics, syntax etc**.** Discourse analysis is among such major areas, with findings that can enhance reading comprehension level among EFL students.Cohesion and coherence in a text are the major topics in discourse analysis which attract many researchers. Investigating cohesiveness of a text is not a new idea, but in association with text structure and textual integrity, it can be a productive topic for research. In the present study the researcher sought to investigate the effect of cohesion, coherence and organizational degree of argumentative texts on Iranian EFL learners' performance in reading comprehension at different proficiency levels.

The Study

There are many different contributive factors which make a text integrative and thus readable. Severalof its aspects have been investigated over the past fifty years. Research has shown that textual integrity can vary in accordance with certain specific text variables and it can either speed or slow reading rates of the texts (Taylor 1990, cited in Weisenmiller, 1999). Textual integrity is of considerable practical significance to educators and publishers of educational materials. For

those who depend upon communication through the printed word, a pertinent issue is whether the materialwill be read and, if read, comprehended by the target readership. The majority of research has focusedon the textual integrity of the text in print. These studies have examined such factors as the effects oftypeface, letter spacing, line spacing, justification contrast, resolution, inverted text, size, type, style,letter spacing, and page layout. These typographic variables have been tested in order to determine various effects upon the reader. Chief among these variables are reading rate and reading comprehension. But more importantly are the contributive factors to discourse structure of a text. "Cohesion and coherence are two important textual elements which are influential on reading a text and understanding it" (Halliday and Hasan 2007; Halliday 2000). Research on cohesion and coherence and their effect on different linguistic traits have been flourishing in recent decades, specifically since the publication of *Cohesion in English* (Halliday and Hasan (2007). Reading comprehension can be viewed as having highly interactive components; namely, the text, reader, comprehension activities, and socio-cultural context. Indeed, empirical studies of readingcomprehension have uncovered some intriguing interactions among text, reader, and task variables (McNamara et al, 1996).

Considering reading as a problematic source of language input, both in academic and non-academic settings, the present study was, therefore, intended to focus on Iranian EFL learners' appreciation of the textual integrity of argumentative texts, and its possible impact on their comprehension at different levels of language proficiency. Thus, the following research questions were addressed and the related hypotheses tested:

RQ1. Does textual integrity of argumentative texts affect male and female Iranian EFL learners'reading comprehension performance?

RQ2. Which level of language proficiency in terms of reading comprehension performance is moreaffected by argumentative texts textual integrity?

Literature Review

The Goal of Reading Comprehension

Kintsch and Rawson (2005) suggest a highly influential theory of comprehension. The theory sees comprehension as depending upon largely automatic processes somewhat akin to the processes

subserving perception. Two major levels of representation are distinguished: a textbase representationthat represents the linguistic structure of the text and its meaning, and a situation model (a mental model of the situation described by the text). The textbase representation will have a number of different levels of representation, including micro-level representations (word and proposition level representations, for example) and a macro-level representation of how ideas in a given passage relateto each other. If this were not complicated enough, for a full understanding, the textbase representation must be related to the situation model, a more abstract representation that is not exclusively verbal and includes a wide range of world knowledge that may include imagery and emotional content.

Perfetti, Landi, and Oakhill (2005) move on to consider how reading comprehension skills develop. They point to the likely critical importance of the learner's ability to identify words fluentlyand retrieve their meanings. In terms of Kintsch and Rawson's model (2005), processes in accuratelyconstructing a textbase representation are critical, and one potential set of limiting factors concerns word identification and access to adequate vocabulary knowledge. Constructing a situation model, however, will require inferences to be made. There are many studies that have attempted to link inferential skills to the development of reading comprehension skills, as well as the development of comprehension monitoring strategies and syntactic skills. A great number of learners have a specificshortfall in developing reading comprehension skills. Innajih (2007) has investigated the effect of explicit instruction of textual markers such as cohesion on the reading comprehension of FL/SL learners. He advises instructors to teach the various types of cohesive devices explicitly andemphasizes their relation to reading comprehension development. Nation (1999) presents a review ofthe nature of the cognitive shortfalls that appear responsible for the problems displayed in these 'poorcomprehenders'. Such learners have weaknesses including the limited extent of vocabulary knowledge as well as higher-level language skills such as inference making. They, therefore, appear to have problems in constructing an adequate text-based representation, though it is possible that theyalso have higher-level problems in constructing a situation model of what they have read.

Argumentative Text structure

Argumentative texts may be organized with different structures. But the most common structure is

asfollows:

1. Introducing the claim.

2.Making the introductory paragraph both interesting and informative. i.e. providing the reader withsufficient background information to be able to understand the claim. For instance, if the claim is about a theory, a brief explanation of the details of the theory should be provided.

3.Presenting the summary of works and key theories being discussed.

4.Supplying the definition of key terms.

5.Giving supporting evidence (reasons, facts, etc.) briefly and coupled with statistics, if necessary, to prove the stated claim and clearly stating how this evidence proves the focused point in the claim.

6.Making the claim very sound and objective and anticipating the readers' objections, i.e. predictingtheir opposing arguments against your arguments; even introducing one or more of the readers' possible anti-arguments and trying to refute them.

7.Concluding the text with the restatement of the claim in a short paragraph and in a different way from what has been stated in the introduction, to show critical thinking, the importance of the claim,and the specific, unambiguous points of the claim.

Research on Textual Integrity

Textual integrity studies aim to find the right fit between the difficulty levels of texts and the readingabilities of students (Ulusoy, 2006). Ulusoy refers to the differences among students in terms of experience and background knowledge about the contents of their course books.

According to McLaughlin (cited in Paula Lissón, 2017), the prediction of textual integrity, which makes a text readable and easy to comprehend, has two main advantages: 1) it helps to decideon the number of people who can be the readers of a special style, and 2) it aids teachers to select appropriate books for their students. Second, it helps authors to understand the extent to which their writings are suitable for their intended readers (p. 69).

Different methods and definitions of textual integrity have been proposed. DuBay (2004) defines textual integrity as what makes a text easier to read. Oosten, Hoste and Tanghe (2011) define it as thedegree of easiness on the part of its addressee to understand its message. Oosten et al. (2011) believethat the concept of textual integrity or unity is subjective in nature, and the

easiness with which a reader can understand a text depends on his or her background knowledge more than anything else.

Quoting Abdollahzadeh and Zolfaghari (2012), to assess textual integrity, Oakland and Lane (2004) introduce three main approaches: qualitative approaches, quantitative approaches, and a combination of these two.

In argumentative texts, the need to persuade through evaluation is central, with a predominance of emotive diction and textual integrity. In such texts, "text forms have a special character, and the ordering must reflect a move from the less to the more evaluative" (Hatim & Mason 1990, p. 193, cited in Jafarinezhad & Tavakoli, 2011). Tirkonnen-Condit (1994, cited in Verzosa Cayago, 2018) views the production of argumentative text as the cognitive process of problem-solving involving thefollowing structural units: situation, problem, solution, and evaluation. Argumentative texts deal withthe mental process of judging. All argumentative texts promote certain beliefs with conceptual relations such as reason, significance, or opposition frequently.

Based on the above-mentioned statements on argumentative texts, the present study addressed the two research questions stated above.

Methodology

Participants

For the purposes of this study, one hundred and twenty Iranian English language learners (male and female), aged between 16 to 22, were selected from a pool of 200 students studying general Englishat Jihad Daneshgahi Language Institute, Isfahan, Iran, to serve as the participants in the research. They were divided into three groups of equal size (40), i.e low proficient, intermediate and high proficient learners based on their scores in an Oxford Quick Placement Test (OQPT). See AppendixA

Design of the Study

This study was conducted through a causal comparative design since the causes of the possible impactof independent variable (textual integrity) on dependent variable (reading comprehension competence) were to be checked. Data was collected from the participants' responses to 6 reading comprehension cloze tests, including 3 textually-intact and 3 textually-distorted texts. Actually,

the selected texts remained either authentic or manipulated, i.e their textual integrity was decreased through deleting certain cohesive devices and disordering paragraphs. The data derived from the threegroups of participants' performances on the tests were then analyzed to obtain results and find answers to the research questions posed.

Instruments and Materials

The instruments and materials used in this study are as follows:

Oxford Quick Placement Test (OQPT)

The first step to collect the required data for the purposes of this study was the administration of a language proficiency test to the participants, i.e OQPT. It was used to divide the participants into different language proficiency levels: low proficient, intermediate and high proficient. According toAllan (2004) this test has been calibrated against the levels system provided by the Common European Framework of Reference for Languages: Learning, Teaching, Assessment (commonly known as the CEF), which has been adopted by the Association of Language Testers in Europe(ALTE) as well as by governments and major institutions, including exam boards, throughout Europe.The version of the test used in this study had 60 multiple-choice questions (See Appendix A) and the participants had to choose the answer from the alternatives for each question. The test evaluated the skills of reading comprehension, vocabulary knowledge, and grammar.

Reading Comprehension Tests: Authentic texts

These cloze tests, which were constructed by the researcher in a random deletion procedure manner,comprised 3 argumentative reading comprehension texts selected from standard reading comprehension pools (i.e. Michigan Test). Efforts were made to choose texts with different subject matter to eliminate the effect of background knowledge on the participants' performance, and with different levels of language difficulty (measured by Flesch Reading Ease, described below) to fit the3 levels of the participants' language proficiency.

Reading Comprehension Tests: Manipulated texts

These 3 cloze tests too were constructed on the same 3 argumentative texts mentioned above.

In devising them, the researcher manipulated the texts in terms of text organization, cohesive devices orties, etc., via a purposeful deletion procedure manner, to become distorted and lose their textual integrity and, therefore, be more difficult to read and understand. Coh-metrix software program wasof great help in this respect.

Coh-metrix Software

Coh-Metrix (developed by McNamara, Louwerse & Graesser, 2002) is a tool used to learn more aboutthe language used in a passage of text. It helps to check how causal, intentional, and other types of connectives help the reader form a more coherent and deeper understanding of the text at the level ofthe causal situation model.

Flesch Reading Ease

Flesch Reading Ease, developed by Flesch (1948, cited in Stone & Parker, 2013) tests the difficulty level of texts. The score on this test will tell you roughly what level of education someone will needto be able to read a piece of text easily. It generates a score usually between 0 and 100. A higher scoremeans the text is easier to read and understand, and a lower score means the text is more difficult forthe reader to read and understand.

Procedure

The following steps were taken to conduct this study: 1) the required participants (120) were selectedfrom a pool of interested English language learners through a proficiency test (OQPT), 2) six readingcomprehension tests (cloze tests - See Appendix B) based on 3 argumentative texts, with different levels of difficulty, were administered (in a counterbalanced manner - a procedure to control the effects of nuisance variables in designs where the same participants are repeatedly subjected to conditions, treatments, or stimuli) at an appropriate time interval to the selected participants, i.e. low proficient, intermediate and high proficient language learners, 3) the participants' performance was scored, 4) the elicited scores were statistically analyzed (using *Paired-Samples t Test and One-Way ANOVA*), and 5) the obtained results were discussed in terms of the posed research questions of thestudy as well as in line with previously-conducted similar research. Table 1 below shows the entire procedure regarding the participants and the tests:

Table 1

Outline of the tests and participants

Sessions	Session 1	Session 2	Session 3 (1 month later)
	OQPT		
Tests		**Cloze Tests** with authentic texts administered to groups 1, 2, 3 (each with 20 participants), and with manipulated texts to groups 4, 5, 6 (each with 20 participants)	**Cloze Tests** with manipulated texts administered to groups 1, 2, 3 (each with 20 participants), and with authentic texts to groups 4, 5, 6 (each with 20 participants)
No. of Participants	200	120	120

Scoring procedure

To score the participants' performance in OQPT, each correct answer was attributed a single point. The total sum of correct answers was used to divide the participants into different English language proficiency levels. To this purpose, the standard deviation and the mean score of their OQPT results were calculated. The participants whose scores were less than one standard deviation below the mean score formed the low proficient group, those with scores falling within the range of one standard deviation above and below the mean formed the intermediate group, and finally, those whose scores were more than one standard deviation above the mean formed the high proficient group.

Data Analysis and Results

The present study aimed to test the following 2 null hypotheses and answer the related research questions:

HO1: Textual integrity of argumentative texts does not affect Iranian EFL Learners' reading

comprehension performance.

HO2: There is no significant difference between the three proficiency groups (i.e low proficient, intermediate, high proficient learners) in terms of the effect of textual integrity on their reading comprehension.

Preliminary Analysis

As outlined above, learners in three proficiency levels (each comprising 40 learners) were recruited for the purpose of the study. These 120 learners were drawn from a pool of 200 EFL learners and were selected based on their scores on the OQPT (Oxford Quick Placement Test). Information regarding the performances of the learners in these three groups on the OQPT is summarized in Table 2.

Table 2

Descriptive Statistics for the Three Groups' OQPT Scores

OQPT	*N*	Minimum	Maximum	Mean	Std. Deviation
Low Proficiency	40	18.00	29.00	24.10	3.07
Intermediate	40	30.00	47.00	39.27	4.65
High Proficiency	40	48.00	54.00	51.20	1.85

It could be seen in Table 2 that on the OQPT, the low proficiency, intermediate, and high proficiencylearners obtained mean scores of 24.10, 39.27, and 51.20 as well as standard deviations of 3.07, 4.65,and 1.85, respectively. The number of learners, the minimum score, and the maximum score in eachgroup are also displayed in the table above. In Table 4.2 below, the results of the normality test (whichis a prerequisite to running the subsequent parametric tests such as *t* test and ANOVA) are presented:

Table 3

Results of the Normality Test

Proficiency Levels	Texts/Tests	Kolmogorov-Smirnov			Shapiro-Wilk		
		Statistic	*Df*	*Sig.*	Statistic	*Df*	*Sig.*
Low Proficiency	Authentic	.12	40	.20	.96	40	.11
	Manipulated	.14	40	.20	.93	40	.08
Intermediate	Authentic	.13	40	.20	.94	40	.10
	Manipulated	.13	40	.20	.95	40	.11
High Proficiency	Authentic	.12	40	.20	.96	40	.12
	Manipulated	.14	40	.20	.94	40	.09

Table 3 shows that for the two tests (i.e., authentic and manipulated) of all the three groups of proficiency, the assumption of normality was met since a *Sig.* value above .05 shows that a distribution was normal, and in this table, all the *Sig.* values lined up under the Kolmogorov-Smirnov test (and under the Shapiro-Wilk test as well) were found to be greater than the significance level of .05. As such, the parametric tests of *t* test and ANOVA could be safely conducted.

Results Effects of Textual Integrity on Reading Comprehension

The first aim of the study was to find out whether textual integrity of argumentative texts had any effects on reading comprehension of such texts by Iranian EFL learners. For this reason, a reading comprehension test with authentic texts and a reading comprehension test with manipulated texts were constructed by the researcher and given to the EFL learners in this study to find out if their comprehension was influenced by textual integrity or not. This was done for all the learners in the three levels of proficiency, considering textual level of difficulty. The results for each proficiency group are presented in separate tables below:

Low Proficiency Learners

To compare the performance of the low proficiency learners on the tests of authentic and manipulated texts, their scores on these two tests were compared by means of a paired-samples *t* test. Table 4 shows the descriptive statistics of this comparison:

Table 4

Results of Descriptive Statistics for Low Proficiency Learners

Level	Texts	*N*	Mean	*Std.* Deviation	*Std.* Error Mean
Low Proficiency	Authentic	40	14.87	1.75	.27
	Manipulated	40	14.02	1.71	.27

The low proficiency learners obtained the mean scores of 14.87 on the tests of authentic texts and their mean score decreased to 14.02 on the test of manipulated texts. To see if the difference between these two mean scores was statistically significant or not, the researcher had to check the paired-samples *t* test table below:

Table 5

Results of Paired-Samples t Test for the Low Proficiency Learners

Paired Differences								
	Mean	*Std.* Deviation	*Std.* Error Mean	95% Confidence Interval of the Difference		*T*	*Df*	*Sig.* (2-tailed)
				Lower	Upper			
Authentic – Manipulated	.85	.73	.11	.61	1.08	7.30	39	.00

The single most important piece of information in Table 5 is the *p* value under the *Sig.* (2-

tailed)column. This value should be compared with the significance level (i.e., .05) to see if the difference between the two sets of scores had been statistically significant or not. A p value less than .05 indicates a significant difference between the two sets of scores, and a p value larger than .05 shows a differencewhich did not reach statistical significance. Since the p value under the *Sig.* (2-tailed) column in Table5 was less than the significance level, it could be inferred that the difference between the authentic ($M = 14.87$) and manipulated ($M = 14.02$) test scores for the low proficiency learners was statisticallysignificant. This is also shown in Figure 1 below:

Figure 1

Mean scores of the low proficiency learners on the two tests

Figure 1 shows that the low proficiency learners' mean score for the authentic reading test was significantly higher than their mean score on the manipulated reading test, giving rise to the conclusion that textual integrity did have an effect on their reading performance.

Intermediate Learners

To compare the performances of the intermediate learners on the two tests, the same statistical

procedure adopted above for low proficiency learners was employed again.

Table 6

Results of Descriptive Statistics for the Intermediate Learners

Level	Texts	N	Mean	Std. Deviation	Std. Error Mean
Intermediate	Authentic	40	16.07	1.04	.16
	Manipulated	40	13.00	1.08	.17

Table 6 indicated that the intermediate learners received the mean score of 16.07 on the test containing authentic texts and the mean score of 13.00 on the test of manipulated texts. To see if this difference between the two mean scores of the intermediate learners was statistically significant or not, the following table had to be checked:

Table 7

Results of Paired-Samples t Test Comparing the Pretest and Posttest Scores of the PCG Learners

Paired Differences								
	Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference		T	Df	Sig. (2-tailed)
				Lower	Upper			
Authentic – Manipulated	3.07	.88	.14	2.79	3.35	21.89	39	.00

Results of paired-samples *t* test in Table 7 revealed that there was a statistically significant difference between the authentic (*M* = 16.07) and manipulated (*M* = 1300) scores of the intermediate learners, *t* (39) = 21.89, *p* = .00. This difference is shown in Figure 4.2 as well:

Figure 2

Mean scores of the intermediate learners on the two tests

Figure 2 above clearly shows that the intermediate learners' performance on the authentic test was considerably better than their performance on the manipulated test, leading to the conclusion that textual integrity did have an impact of their reading performance in their L2.

High Proficiency Learners

The results for the comparison of the scores obtained from the two tests taken by the high proficiencylearners are displayed below:

Table 8

Results of Descriptive Statistics for the High Proficiency Learners

Level	Texts	*N*	Mean	*Std.* Deviation	*Std.* Error Mean
Intermediate	Authentic	40	16.35	.97	.15
	Manipulated	40	15.40	.67	.10

Table 8 showed that the high proficiency learners' mean score was 16.35 on the authentic texts test and 15.40 on the manipulated texts test. To figure out whether this difference between the two meansscores of the high proficiency learners was statistically significant or not, Table 4.8 had

to be consulted:

Table 9

Results of Paired-Samples t Test Comparing the High Proficiency Learners

Paired Differences								
	Mean	*Std.* Deviation	*Std.* Err orMean	95% Confidence Interval of the Difference		*t*	*Df*	*Sig.* (2-tailed)
				Lower	Upper			
Authentic – Manipulated	.95	.71	.11	.72	1.17	8.41	39	.00

Results of paired-samples *t* test in Table 4.8 demonstrated that there was a statistically significant difference between the authentic (*M* = 16.35) and manipulated (*M* = 15.40) test scores ofthe high proficiency learners, *t* (39) = 8.41, *p* = .00. This difference is also represented in Figure 4.3 below:

Figure 3

Mean scores of the high proficiency learners *on the two tests*

Figure 4.3 clearly shows that the high proficiency learners' performance was superior on the test of authentic texts, compared to their performance on the test of manipulated texts, indicating thattextual integrity influenced the reading performance of high proficiency EFL learners.

Comparing the Three Levels of Proficiency

Another objective of the study was to find out whether textual integrity affected the learners at different proficiency levels equally or not. For this purpose, for the learners in each proficiency level, a difference score was computed by subtracting the manipulated test scores from the authentic test scores. This way, three sets of difference scores (for the learners in the three different proficiency groups) were obtained. Then, one-way between-groups ANOVA was employed to compare these three sets of scores.

Table 10

Descriptive Statistics Results Comparing Learners' Difference Scores

| | | | | | 95% Confidence Interval for Mean | | | |
	N	Mean	*Std.* Deviation	*Std.* Error	Lower Bound	Upper Bound	Minimum	Maximum
Low	40	.85	.73	.11	.61	1.08	.00	2.00

Proficiency								
Intermediate	40	3.07	.88	.14	2.79	3.35	1.00	5.00
High Proficiency	40	.95	.71	.11	.72	1.17	.00	2.00
Total	120	1.62	1.29	.11	1.39	1.85	.00	5.00

The mean difference scores of the low proficiency ($M = .85$), intermediate ($M = 3.07$), and high proficiency ($M = .95$) learners are shown in Table 4.9 above. This shows that the difference betweenauthentic and manipulated scores was highest for intermediate learners and lowest for low proficiency learners. To figure out whether the differences among these mean scores were significant or not, oneneeded to check the *p* value under the *Sig.* column in the ANOVA table below:

Table 11

Results of One-Way ANOVA for Comparing the Learners' Difference Scores

	Sum of Squares	Df	Mean Square	F	Sig.
Between Groups	126.35	2	63.17	102.98	.00
Within Groups	71.77	117	.61		
Total	198.12	119			

As is displayed in Table 11, there was a statistically significant difference in the difference scores for low proficiency ($M = .85$), intermediate ($M = 3.07$), and high proficiency ($M = .95$) learnersbecause the *p* value under the *Sig.* column was lower than the specified level of significance (i.e., .00 < .05). To pinpoint the exact location of the difference(s), the post hoc test table (Table 12) had to bechecked:

Table 12

Scheffe Post Hoc Test Results for Comparing Learners' Difference Scores

Lavenezhad: Effect of Textual Integrity of Argumentative Texts on EFL Learners' Reading Performance

Groups		Mean Difference	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Low	Intermediate	-2.22*	.17	.00	-2.65	-1.79
Low	High	-.10	.17	.85	-.53	.33
Intermediate	High	2.12*	.17	.00	1.69	2.55

Based on the information presented in Table 12, the difference between low proficiency (*M* =.85) and intermediate (*M* = 3.07) learners was statistically significant ($p < .05$), but the difference between low proficiency and high proficiency learners (*M* = .95) failed to reach statistical significance. Besides, there was a significant difference between intermediate and high proficiency learners. All this implies that the effect of textual integrity on the comprehension of argumentative texts was significantly obvious for intermediate learners, but less obvious for low proficiency and high proficiency learners. This result is also graphically represented through the bar graph in Figure 4:

Figure 4

The mean difference scores for the three groups of proficiency

The bar graph in Figure 4 shows that intermediate learners had the highest difference score compared to the low difference scores of the low proficiency and high proficiency learners. This implies that textual integrity had the highest effect on the comprehension of argumentative texts by intermediate EFL learners.

Discussion

The first research question of study was: "Does textual integrity of argumentative texts affect IranianEFL Learners' reading comprehension performance?" In line with this question, the first nullhypothesis posited: "Textual integrity of argumentative texts does not affect Iranian EFL Learners' reading comprehension performance." In order to investigate this issue among the three levels of proficiency, three separate paired-samples *t*-tests were run. The results of the data analysis presentedin Chapter 4 indicated that for all the three levels of proficiency there was a statistically significant difference between the authentic and manipulated test scores of the EFL learners. In fact, the first nullhypothesis of the study was rejected, showing that textual integrity of argumentative texts has a significant effect on Iranian EFL learners' reading comprehension performance. This finding is in line with that of Innajih (2007) who investigated the effect of explicit instruction of textual markers on the reading comprehension of FL/SL learners. In line with our finding, he demanded that instructorsteach the various types of cohesive devices explicitly and emphasized their relation to reading comprehension. Also, in a similar study involving cloze tests, Smith (2004), found a significant relationship between reading comprehension and the comprehension of discourse markers.

Moreover, stressing the importance of textual integrity for reading comprehension, Moradan (1995) suggested that explicit instruction of connectors and linking words should be involved in language courses to help learners take advantage of their knowledge of them in reading comprehension and other language uses. Another study that lends further support to our finding is thestudy by Basaraba et al. (2011). In their study, they claim that reading passages that are rich in textual integrity or employment of appropriate discourse connectors help free up additional cognitiveresources that can be applied by EFL learners to understand the meaning of words, phrases, and sentences within text, and thus making L2 comprehension happen more easily and rapidly.

The positive impacts of knowing and using the markers of textual integrity have also been reported on other skills. For instance, Aidinlou (2012), Emmanuel (2013) and Jalilifar (2008) reported that cohesive devices are fundamental linguistic devices which lead the readers to the direction of the flow of text. In general, these studies also conclude that there is a significant relationship between the higher use of cohesive devices and the quality of students' written production. Nevertheless, the findings of the present study run counter to the findings of Castro, (2004) who concluded that the use of cohesive devices and quality of writing are not soundly consistent.

The second research question of the study was: "Which level of language proficiency in terms of reading comprehension performance is more affected by argumentative texts' textual integrity?" In line with this question, the second null hypothesis argued: "There is no significant difference between the three proficiency groups (i.e low proficient, intermediate, high proficient learners) in terms of the effect of textual integrity on their reading comprehension." In order to examine this second hypothesis, three sets of difference scores (for the learners in the three different proficiency groups) were obtained through the procedures stated in Chapter 3, and one-way between-groups ANOVA was employed to compare these three sets of scores.

The results of ANOVA showed that the difference between authentic and manipulated scores was highest for intermediate learners and lowest for low proficiency learners. Moreover, the difference between low proficiency ($M = .85$) and intermediate ($M = 3.07$) learners was statistically significant ($p < .05$), but the difference between low proficiency and high proficiency learners ($M =.95$) failed to reach statistical significance. Besides, there was a significant difference between intermediate and high proficiency learners. All this leads us to reject the second hull hypothesis of the study, claiming that the effect of textual integrity on the comprehension of argumentative texts was significantly obvious for intermediate learners, but less obvious for low proficiency and high proficiency learners. These findings lend further support to the findings of other researchers who have emphasized the importance of different instructional activities including the teaching of textual integrity markers to intermediate EFL learners (Block & Pressley, 2002; Duke & Pearson, 2002). These studies frequently present isolated instructional practices that influence intermediate L2 students' reading achievement.

It can be claimed based on previous literature that reading represents a form of thinking

(Kurland, 2000; Paul, 1995), and accordingly researchers argue that certain levels of reading comprehension require acts of cognition, such as analysis, synthesis, and interpretation (Roe, Smith,& Burns, 2005). One justification for the findings related to our second research question might be attributed to the fact that learners with an intermediate level proficiency in this study benefited fromthese thinking resources more than those with low and advanced levels of English proficiency.

Nevertheless, there are other researchers (Basaraba et al., 2011) who have claimed that reading comprehension improves when student thinking abilities improve. In other words, improved cognition enables improved comprehension. This is opposed to our findings because based on such claims we expected the learners in the advanced group to benefit most from textual integrity for betterreading comprehension performance.

Conclusions

One of the objectives of the present study was to examine the effect of textual integrity of argumentative texts on Iranian EFL learners' reading comprehension performance. Another objectivewas to determine whether there is a significant difference between the three proficiency groups (i.e low proficient, intermediate, high proficient learners) in terms of the effect of textual integrity on their reading comprehension performance.

The analysis of data resulted in two major conclusions drawn from the study. Firstly, the reading comprehension of all the three groups of proficiency were significantly influenced by textual integrity; in other words, it was found that the texts that were authentic in terms of textual integrity (i.e. text organization, cohesive devices, etc.) were easier to comprehend for all the three groups at different levels of proficiency, compared to those texts which were manipulated to lose their textual unity. Secondly, textual integrity had the highest effect on the comprehension performance of intermediate learners rather than low proficiency or high proficiency learners. This fact demonstratedthat learners at the intermediate level proficiency in this study benefited from textual integrity for higher L2 comprehension more than low and advanced proficiency levels.

The findings of the present study have implications for EFL learners, teachers, and materials developers in the realm of FL and SL teaching in particular and education in general. Moreover, the findings of this study could enrich the literature in the area of foreign language learning,

especially Iranian EFL learners' reading comprehension performance. Also, the findings of the study can be used by language practitioners and curriculum developers to consider students' needs for reading passages that are rich in markers of textual integrity. In fact, in choosing instructional materials for reading comprehension classes, passages which enjoy higher levels of textual integrity can be used to create a better educational context in which EFL learners' reading competence can be developed. Based upon this fact, syllabus designers should realize that including these elements, i.e cohesive tiesand/or connectors as well as text organization, in textbooks and materials is essential.

Another implication which can be drawn from the study is to require teachers to explain thoroughly certain aspects of textual integrity and cohesive devices for their delicacy and subtlety (substitution reference, some adversative conjunctions, etc.). As it is evident that the learners' readingcomprehension improved when they received texts that were authentic in terms of textual integrity, it is suggested that the teaching and use of such elements be intensified in-class and out-of-class alongside the assignment of reading passages to motivate students for extensive reading.

References

Abdollahzadeh, E., & Zolfaghari (2012). Qualitative and Quantitative Examination of Text Type Readabilities: A Comparative Analysis, *RALS, 3(2),*3-4.

Aidinlou, N. (2012). The effect of discourse markers ınstruction on EFL learners' writing. *World Journal of Education. 2(2),* 10-16.

Anderson, T. H. (2001). Study skills and Learning Strategies. Theoretical Report, 104. University of Illinois at Urbana-Champaign.

Azer, M. (1998), Argumentative Text – Its Substance and Characteristics, Lecture [in Hebrew] at the Israeli Applied Linguistics Conference at the Haifa University.

Basaraba, D., Travers, P., & Chaparro, E. (2011). Application of Ehri's theory: Instructional implications of students' decoding skills. Paper presented at the National Association of SchoolPsychology annual conference, San Diego, CA.

Block, C. C., & Pressley, M. (Eds.). (2002). *Comprehension instruction: Research-based best practices.* New York: Guilford Press.

Bugel, K., & Bunk, B. P. (1990). Six Differences in Foreign Language Text Comprehension: The

roleof interest and prior knowledge. *Modern Language Journal*, *80 (1),*17.

Castro, C. (2004). Cohesion and the social construction of meaning in the essays of Filipino collegestudents writing in L2 English. *Asia Pacific Education Review, 5(2)*, 215-225.

Chastain, K. (1971). *The Development of Modern Language Skills: Theory to Practice.* Philadelphia,Penn.: Curriculum Development Center.

DuBay, W. H. (2004). *The principles of readability*. California: Impact Information.

Duke, N. K., & Pearson, P. D. (2002). Effective practices for developing reading comprehension. In A. E. Farstrup & S. Samuels (Eds.), *What research has to say about reading instruction* (pp.205-242). Newark, DE: International Reading Association.

Emmanuel, C. (2013). An analysis of discourse markers ın academic report writing: pedagogical ımplications. *International Journal of Academic Research And Reflection, 1(3)*, 15-24.

Halliday, M.A.K. (2000). Introduction to Functional Grammar, (2ⁿᵈ ed.). Foreign Language TeachingResearch Press, Beijing.

Halliday, M.A.K., & Hasan, R. (2007). Cohesion in English: Fifth Edition. Longman: London. Hatim, B., & Mason, I. (1990). Discourse and the translator. New York: Longman Inc.

Innajih, A. (2007). "The effect of conjunctive types on the English language reading comprehension of Libyan university students."   [Online]   Available:www.ecls.ac.Uk /publish/text/the Effect of Conjunctive/

Jafarinejad, R., Tavakoli, M. (2011). Investigating the relationship between discourse markers,language proficiency and reading comprehension: A case of some Iranian university students. *Procedia Social and Behavioral Sciences*, 1526–1530.

Jalilifar, A. (2008). Discourse markers in composition writings: The case of Iranian learners of Englishas a foreign language. *English Language Teaching. 1(2),* 114-122.

Kinneavy, J. (1971), A Theory of Discourse, New-Jersey, Prentice-Hall, Inc, Englewood Cliffs. Kintsch, W. & Rawson, D. (2005). *The Acquisition of Reading Comprehension Skill*. In M. J. Snowling, & C. Hulme (Eds.), The Science of Reading: A Handbook (pp. 227-247). Oxford: Blackwell.

Kurland, D. J. (2000). Critical reading v. critical thinking. In *How language really works: The fundamentals of critical reading and effective writing*. Retrieved August 30, 2005, from

http://criticalreading.com/critical_reading_thinking.htm

Lissón, P. (2017). Investigating the use of readability metrics to detect differences in written productions of learners: A corpus-based study. Bellaterra *Journal of Teaching & Learning Language & Literature, 10 (4)*, p. 69.

McNamara, D. S., Louwerse, M. M., & Graesser, A. C. (2002). Coh-Metrix (Version 2.0) [Software]. Memphis, TN: University of Memphis, Institute for Intelligent Systems. Available from http://cohmetrix.memphis.edu/cohmetrixpr/index.html.

Moradan, A. (1995). *Significance of conjunctions as a cohesive device in teaching writing.* Unpublished MA thesis. Allameh Tabatabai University.

Oosten, V. P, Hoste, E., & Tanghe, D. (2011). A posteriori agreement as a quality measure forreadability prediction systems. *CICLing*, 2, 424-435.

Paul, R. (1995). *Critical thinking: How to prepare students for a rapidly changing world*. Santa Rosa,CA: The Foundation for Critical Thinking.

Pearson, D., P. (2002). Handbook of Reading Research. Lawrence Erlbaum Associates Publishers,London.

Perfetti, C. A., Landi, N., & Oakhill, J. (2005). *The Acquisition of Reading Comprehension Skill*. In M. J. Snowling, & C. Hulme (Eds.), The Science of Reading: A Handbook (227-247). Oxford:Blackwell.

Rivers, W. M. (1968). *Teaching Foreign Language Skills*. University of Chicago Press.

Roe, B., Smith, S. H., & Burns, P. C. (2005). *Teaching reading in today's elementary schools* (9th ed.). Boston: Houghton Mifflin.

Smith, F. (2004). *Understanding reading: A psycholinguistic analysis of reading and learning to read*. New York: Holt, Rinehart and Winston.

Stone, G. & Parker, L. (2013) Developing the Flesch reading ease formula for the contemporaryaccounting communications landscape. *Qualitative Research in Accounting & Management,* 10(1), 31-59.

Verzosa Cayago, A. (2018). Gender-Preferential Use of Rhetorical Structure and Metadiscourse Markers in Argumentative Text. International Conference on Research and Publication, Clark,Philippines.

Ulusoy, M. (2006). Readability approaches: Implications for Turkey. *International Education*

Journal,7(3), 323-332.

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

Maryam Asgari [1*]

[1]English Department, Islamic Azad University, Bandar Abbas, Iran

Abstract

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'.

[1]Corresponding Author's Email:
 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

Negative effect on words in St	Penalty for negative effect
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)

	M	*SD*
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) where checked. Figures 1, 2, and 3 belowdepict 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)

Asgari: Relationship between Two Translation Quality Assessments: Holistic Rating and Waddington's ...

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

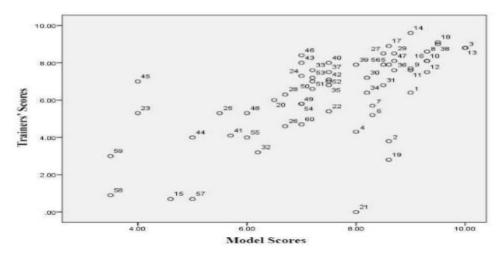   Before testing our research hypothesis, it is necessary to check for the correlation between overall scores obtained by assessing papers using Waddington's model by three raters and the scores obtained by assessing papers using Waddington's model by the researcher herself. However, in this case, Pearson's 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

	Overall Rating Scores	
Model Scores	*rho*	*P*
	.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	
	rho	*P*
	.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 (α=.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 [Accessed 18 Jan 2016].

House, J. (2001). Translation Quality Assessment: Linguistic Description versus Social Evaluation. *Meta: Translator's Journal. 46*(2).

Khanmohammd, 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. [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.

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

Maryam Asgari [1*]

[1]English Department, Islamic Azad University, Bandar Abbas, Iran

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'.

[1]Corresponding Author's Email:
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.*

<div align="center">

Method

</div>

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

Negative effect on words in St	Penalty for negative effect
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)

	M	*SD*
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) where checked. Figures 1, 2, and 3 belowdepict histograms for instructors' scores, model scores and rating scores respectively.

Figure 1

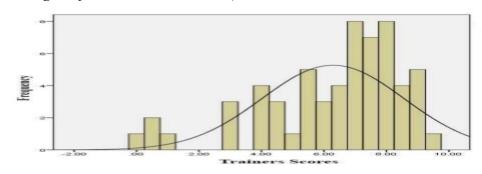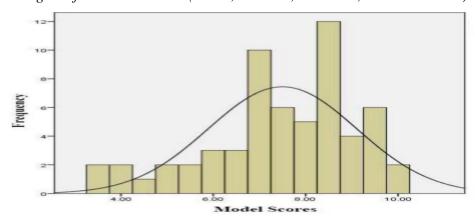*Histogram for Instructors' Scores (N=60, M=6.29, SD=2.27, Skewness=-1.05)*

Figure 2

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

Asgari: Relationship between Two Translation Quality Assessments: Holistic Rating and Waddington's ...
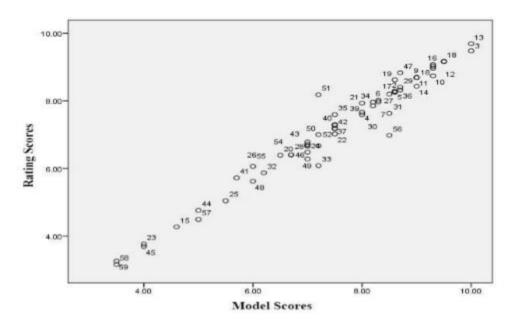
All figures above show that scores are negatively skewed. Skewness values of minus 1.05, minus .82 and minus .80 show that, to some extent, the assumption of normality has been violated. In order to look for possible violations of the assumption of linearity and to locate outliers, scatterplots for the 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.
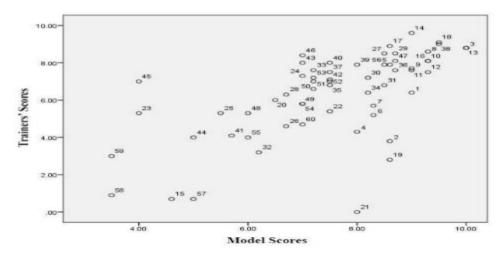
Figure 3

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

	Overall Rating Scores	
Model Scores	*rho*	*P*
	.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	
	rho	*P*
	.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 (α=.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 [Accessed 18 Jan 2016].

House, J. (2001). Translation Quality Assessment: Linguistic Description versus Social Evaluation. *Meta: Translator's Journal. 46*(2).

Khanmohammd, 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. [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.

Impact of Opinion-Exchange and Information Gap Tasks on EFL Learners' Willingness to Communicate

Mahsa Soleimani[1] & Hossein Vahid Dastjerdi[2*]

[1,2]English Department, University of Isfahan, Isfahan, Iran

Abstract

Available online

Keywords:

Willingness to Communicate, information-gap, task, opinion-exchange task

The present study sought to investigate the effect of two types of tasks; namely, 'information-gap' and 'opinion-exchange', on EFL learners' Willingness to Communicate (WTC). To this end, the needed data was collected from 90 female EFL learners in Koushesh Language Institute in Isfahan, Iran, and the participants were divided into one control and two experimental groups (A and B). The latter received treatments in one of the two experimental situations and were measured for their level of L2 Willingness to Communicate (WTC). Group A received 'information gap' for a period of 16 sessions, and group B received 'opinion-exchange' tasks. The control group received no task of specific type. At the end of the treatment, the participants filled a questionnaire on their WTC. The results indicated that members of both the experimental groups outperformed those of the control group. Furthermore, it was revealed that opinion-exchange tasks had better effects on the enhancement of the participants' WTC. The findings of this study may be very beneficial for the teachers of the English language who wish to improve their EFL learners' speaking ability. In fact, creating environments for learners to communicate in English inside and outside the classroom through tasks would enhance learners' willingness to communicate.

[2]Corresponding Author's Email:
h_vahid@yahoo.com

Introduction

Task-based language teaching (TBLT) is one of the recent language teaching methodsthat aims to improve learners' speaking ability by engaging them in different tasks (Richards & Rogers, 2001). A task-based approach attempts to put the learners in the natural context needed for language use. As learners endeavor to perform a task, they will have opportunities to interact and to use their L2. Such an interaction is thoughtto make language acquisition easier for learners to work, to understand, and to expressthemselves. By doing so, they check to see if they have understood the message correctly and, sometimes, they have to ask for clarification. Through interaction with others, language learners listen to the language that may be beyond their present ability, but may be integrated into their knowledge of L2 for use in the future. (Ellis, 2009), ''the central purpose we are concerned with is language learning, and task presents this in the form of a negotiation between knowledge that the learner holdsand new knowledge''. Task-based language learning (TBLL) does not include specification of a sequence for language teaching but consists of a sequence ofcommunicative tasks to be done in the L2 (Nunan, 2005). It gives learners a sample of the target language which is organized based on the purposes which people uselanguage (Richards & Renandya, 2002, p. 94).

The literature on TBLT has witnessed several classifications of tasks based on factors such as time, mental or linguistic resources needed, or their being real-life or pedagogical (Ellis, 2008). An interesting line of inquiry has been to investigate the potential of different tasks to foster learners' speaking ability (e.g., Ellis, 2009; Murad, 2009).

Task-based teaching is a generally approach to teaching process. The teaching method is seen as a set of communicative tasks that are directly related to curricular aims. Nunan (1991) views the task as a part of meaning-focused work, which involves learners in comprehending, manipulating, producing and interacting in the targetlanguage.

There are a number of different explanations in the literature about what a task really is. However, many researchers today make an important difference between target tasks, which students need to do outside the classroom, and pedagogical tasks, which form the base of the classroom activity during the teaching.

As far as target tasks are concerned, Song and Zhang (2008) list a number of them for example: buying a pair of shoes, making an airline reservation, borrowing a library book, taking a driving test, typing a letter and making a hotel reservation. He sums up a description of task in this way: by task, it is meant the hundred and one things people do in everyday life, at work, at play, and in –between. Tasks are things that students typically do outside of the classroom. The last rationale for language instruction is to enable learners to complete these activities successfully in the real world using the target language.

Willingness to Communicate (WTC) and Second LanguageAcquisition (SLA)

As more attention is paid to meaningful communication, more attention would be paid to Willingness to Communicate (WTC, hereafter). This is why investigation of WTC has always been the focus of those researchers overwhelmed with meaningful communication. This is due to the notion that communication has been replaced with other aims of language learning in modern world of L2 learning and teaching and, in this way, 'communication' stands for 'meaningful communication'. Therefore, when communication is deemed as the primary objective of language learning, WTC gets more and more significant for EFL learners.

Despite the vital role of WTC in rendering meaningful communication that is boosted in EFL contexts, only little attention has been paid to it. In this vein, Akbarzadeh and Narafshan (2016) put that WTC is a variable that has been frequentlyforegrounded in literature. They continue that for learning to talk in the L2, learners need to be willing to communicate in the L2. Therefore, developing learners` L2 WTC should be the fundamental goal of language instruction. Moreover, MacCIntyre et al. (2001) define WTC as a readiness to enter the discourse at a particular time with a specific person or persons. Considering WTC as a situational construct, researchers have examined how it is influenced by situational variables such as contextual variables and social support. As a conclusion to this, in a student-friendly environment student would be more willing to talk in class than a teacher-centered class. In a stress-free supporting environment, learners can build a better rapport not only with each other but also with the teacher, which will in turn boost the learning process to a considerable extent. Then willingness to communicate is very important in Second Language Acquisition. Second language acquisition, or SLA, has twomeanings. In a general sense, it is a term to define learning a second language. Specially, it is the name of the

theory of the procedure by which we acquire - or pick up - a second language. This is mainly a subconscious process, which occurs while wefocus on communication. It can be related with second language learning, which describes how formal language teaching helps us learn language through more conscious processes.

Literature Review

Many L2 teachers consider students' silence in the TESOL profession as a negative attribute. Evidence has shown that L2 teachers find it as a problem to get students to answer willingly in classrooms. It appears that learners whodo not participate in L2 interaction are usually regarded as being passive and unmotivated.

In the TESOL profession, which was dominated by communicative language teaching (CLT) methodology since the1970, researchers' common goal was to motivate learners to become communicators that are more effective. As Breen (2001) observes, "one of CLT's innovations was to advocate spontaneous learner communication through talk about topics and issues that were immediatelymeaningful to them" (p. 113). It is undeniable that students' participation is very important in language learning. Interaction research, for instance, provides abundant evidence for the simplifying role that participation plays in language acquisition. Therefore, researchers such as Macintyre, Dornyei, Clement, and Noels (1998) argue for the importance of promoting learners' WTC in L2 education.

English, which is defined as an international language, is used by more than one and ahalf billion people (Strevens, 1992) as a first, second, or foreign language for communication purposes. Therefore, the purpose of teaching English has moved from the mastery of structure to the ability to use the language for communicative purposes. Jahanshahi (2013) investigated whether college students who were learning English as a foreign language in the Iranian context were willing to communicate when they had an opportunity. The study utilized a quantitative data collection and analysis procedures. The appropriate method was quasi experimental design and data were analyzed through t-test. The total participants in this study were 80 English students (either translation or teaching) in South Tehran branch of Islamic Azad University. The Willingness to

Communicate questionnaire was distributed among theparticipants before and after the treatment (Group Work). The results showed that Group activities had a important effect on students' willingness to communicate and revealed that learners were willing to communicate in English. Participants preferred using group work; in such relax and calm atmosphere and enjoyable way to solve problems and go to classes they were more willing to communicate. Creating environments for learners to communicate in English inside and outside the classroom and via internet and synchronous chat would enhance learners willing to communicate.

Foroutanfar (2015) explored and compared the effectiveness of focus-on-form (FonF) task instruction (i.e., input vs. collaborative output tasks) on Iranian English asa foreign language (EFL) learners' willingness to communicate (WTC). To achieve the objective, 50 Iranian pre intermediate EFL students took part in this study. The FonF tasks were presented in the form of input FonF tasks (i.e., textual enhancement, processing instruction, and discourse tasks) and collaborative output FonF tasks (i.e., dictoglass, jigsaw, and text reconstruction tasks) to 2 classes of 50 pre-intermediate Iranian EFL students. Participants answered the pre-test-post-test questions of the WTCquestionnaire. The analysis of the paired samples t test in the input and collaborative output FonF tasks group showed that these types of FonF tasks did not improve the participants' WTC significantly. However, the participants in the collaborative output FonFtasks group outperformed those in the input FonF tasks groups. The results of thestudy imply that monitoring students' WTC in second/foreign language (L2) and improving it should be considered as one of the goals of L2 teachers and syllabus designers through implementation of collaborative FonF tasks.

Razmjoo (2015) investigated the effect of competitive and cooperative teaching on Iranian EFL learners' WTC. The participants included 120 Iranian femaleintermediate students studying at Ayandegan high school, in Shiraz. A 27-item likert type WTC questionnaire developed by MacIntyre et al (2001) was administered as a pre-test, then the learners were divided into two groups randomly, one group experienced the competitive teaching and the other group experienced the cooperativeteaching. The treatment took 20 sessions. The Fundamentals A from Top Notch Serieswas used as the material taught to learners. At the end of the experiment, the same WTC questionnaire was once again administered as the post-test. A paired sample t- Test was run on the mean scores of both groups to find the effect of teachingapproaches on learners' WTC. The results

showed that the cooperative teaching had a significant effect on learners' WTC. This means that different methods of teaching areinfluential on WTC of Iranian EFL learners.

One of the pedagogical implications of the research on the Willingness to Communicate (WTC) might be to propose practical ways of making language learnersmore willing to communicate in the classroom. Abdollahzadeh (2014) investigated the impact of teaching communication strategies (CSs) on Iranian EFL learners' WTC.To this end, 8 intact classes were included as the experimental and control groups. The control group underwent regular language instruction, while the experimental group received the treatment (i.e., communication strategy training). The self-report measurement of WTC (MacIntyre, Baker, Clément, & Conrad, 2001) was done before(pre-test) and after the treatment (post-test). The results of the independent-samples t- test showed that the degree of WTC of the treatment group was significantly higher compared with that of the control group. It was concluded that teaching CSs helps learners become more willing to communicate in the classroom.

Fatemipour and Mohammadi (2014) investigated different information-gap activities, and their effect on learners' willingness to communicate. More specifically,this study examined the effect of jigsaw, missing-information, and finding the differences activities on the learners' willingness to communicate. Using non- probability sampling, KET (Key English Test) was administered as test of homogeneity to select the research subjects. A total of 60 participants above the ageof 13 who were at the elementary level participated in this study. The subjects were assigned to three experimental groups. A questionnaire was also administered tomeasure the learners' willingness to communicate with the WTC scale, as a pre-test at the first session of the term. The treatment included the application of the three information-gap activities, in the three experimental groups, and at the end, the same questionnaire was applied as the post-test. The obtained results showed that therewere significant differences regarding the effects of information-gap activities on

Previous studies have revealed that there is a positive relationship between task-based language teaching and interaction between students. Although the term willingness to communicate may not seem a brand- new concept to be studied; however, the context in which it was implemented seems different in TBLT since willingness to communicative is enriched contextually. Task-based was

thought to improve the atmosphere for learners to practice social contexts of the communication. Consequently, this study intended to determine to what extent two different tasks (opinion exchange task and information gap task) based on task-based languageteaching would influence EFL learners' speaking ability and willingness to communicate. Moreover, it sought to investigate whether these would be able to improve (any or both of) these two variables or not. (Maftoon & Sarem, 2013).

Several researchers have tried to examine ways in which learners' willingness to communicate could be fostered so that they can use their language for communication (Zarrinabadi, 2014). These studies have found several psychological and linguistic variables that influence learners' willingness to communicate (Ellis, 2009). An interesting line of inquiry has been to investigate the effect of teacher activities on promoting learners' willingness to communicate. Past research has shown that tasks are useful ways to enhance the students' speaking ability. Thepresent study aimed at applying information-gap and opinion-exchange tasks forIranian EFL learners to explore their effectiveness in improving their willingness to communicate skills.

The objective of this study was examining the effect of information-gap and opinion- exchange tasks on Iranian EFL learners' willingness to communicate. This studyactually aimed to find out whether different task types have the capacity to improve learners' willingness to communicate in the Iranian EFL context. Thus, the following research questions were addressed:

RQ1: Does information-gap task have a significant effect on Iranian EFL learners' willingness to communicate?

RQ2: Does opinion-exchange task have a significant effect on Iranian EFL learners' willingness to communicate?

RQ3: Is there a significant difference between the effects of information-gap and opinion-exchange tasks on WTC?

Methodology

Design of the Study

An explanatory sequential mixed-methods design was used to collect and analyze the data of the study, and to explore the impact of opinion-exchange and information-gap tasks on learners' L2

WTC.

Participants

A sample of 90 female language learners of 13 to 18 years of age, learning English as a Foreign Language in Koushesh Language Institute of Isfahan, were selected based on OPT as the participants of this study. They were homogenized in terms of language proficiency.

Instruments

Oxford Placement Test

The Oxford Placement Test (OPT) was used to measure the participants' language proficiency. In fact, OPT was used to make sure about the participants' homogeneity prior to the application of the treatment.

Semi-Structured Interviews

Semi-structured interviews were used in this study as qualitative tools for data collection.

Willingness to Communicate Questionnaire

A self-report questionnaire was used for the aim of data collection at the quantitative phase of the study. The questionnaire was developed by MacIntyre et al. (2001). It comprises 26 items related to willingness to communicate inside the classroom. It has items assessing willingness to communicate in four language skills: reading, writing, speaking and comprehension. It is a Likert-type questionnaire ranging from 1 to 5, in which 1 is 'almost never willing' and 5 is 'almost always willing'.

Procedure

Data Collection Procedure

Oxford Placement Test (OPT) was first implemented for the selection of the homogeneous participants. The selected participants were randomly assigned to experimental and control groups.

Moreover, they were homogenized in terms of their willingness to communicate using their scores on a questionnaire developed by (MacIntyre et al., 2001).

In the next step, L2 learners' received treatments in one of the two experimental situations and were measured for their level of L2 WTC. Here, group A received information gap for a period of 16 sessions, while group B received opinion- exchange tasks. The control group received no task of specific type. At the end of the treatment, the participants filled a questionnaire on their WTC.

The experimental results, therefore, guided the development of a second, qualitative phase. The researcher developed and refined the qualitative research questions and implemented the qualitative phase in which semi-structured interview data were gathered and analyzed to assist the explanation and interpretation of the quantitative results. In the interviews, the participants were asked to talk their feelingstoward speaking in the classroom and the reasons for their feelings. Meanwhile,whenever the students referred to activities in the classroom, some probing questions were used.

Data Analysis Procedure

As mentioned above, there were three groups in the study and one dependent variable (WTC). Thus, a one-way analysis of variance (ANOVA) was usedto analyze the results of the quantitative phase. Then, Tukey post hoc test was used to examine the areas of difference between the groups. For the qualitative data, the qualitative content analysis by Miles and Huberman (1994) was used to identify the major themes and reasons behind any possible effect of the two types of tasks on learners' willingness to communicate.

Results

The analysis of the data entailed the following tabulated results:

Table 1

T-test for the Comparison of Homogeneity

Levene's Test for Equalityof Variances		t-test for Equality of Mans				
	F	Sig.	t	df	Sig. (2-tailed)	95% Confidence Interval of the Difference

							Lower	Upper
VAR00001	Equal variances assumed	.014	.906	.552	38	.584	-2.10101	1.20101
	Equal variancesnot assumed			-.552	37.984	.584	-2.10103	1.20103

According to the above table, there is no significant difference between the performances of the participants in the OPT (Sig. = 0.584 > 0.05). So, the participantswere all at the same level of proficiency prior to the treatment.

In order to investigate the effect of using information gap task oncommunicative willingness, two sets of one-way ANOVA were run. The first one was between the participants' scores of the three groups on pre-test and the second one between immediate post-tests of three groups. The descriptive statistics of pre-test are presented in Table 2 below.

Table 2

Descriptive Statistics of the Participants' Performance on Pre-test

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
Prettest of information gapgroup	30	1.3333	.72375	.18687	.9325	1.7341	.00	2.00
Prettest of opinion exchange tasks group	30	1.4000	.73679	.19024	.9920	1.8080	.00	3.00
Prettest of Control group	30	1.1333	.74322	.19190	.7217	1.5449	.00	2.00

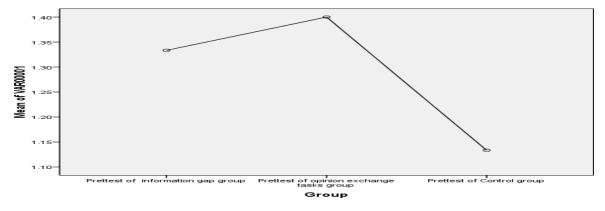| Total | 90 1.2889 | .72683 | .10835 | 1.0705 | 1.5073 | .00 3.00 |

According to Table 2, means of the three groups were almost the same. After collecting the data, ANOVA was performed using SPSS to see whether the differences between the three groups were significant or not. The results are presentedin Table 3 below.

Table 3

One-way ANOVA for Comparing the Performance of Groups

	Sum of Squares	*df*	Mean Square	*F*	*Sig.*
Between Groups	.578	2	.289	.535	.589
Within Groups	22.667	42	.540		
Total	23.244	44			

According to Table 3, the mean differences between the three groups were notsignificant (Sig= 0.589). This shows that the participants in the three groups were at the same level of knowledge. Figure 1 below shows the graphical comparison of the two groups on the pre-test.

Figure 1

Graphical Representation of the Groups' Performance on the Pre-test

To illuminate where the significant differences exist among the groups, post hoc test (with an

alpha level of .05) was conducted. The results are shown in Table4.

Table 4

Post Hoc and Multiple Comparison of Three Groups

(I) Group	(J) Group	Mean Difference(I-J)	Std. Error	Sig.	95% Confidence Interval Lower Bound	Upper Bound
Prettest of information gap group	Prettest of opinion exchange tasks group	-.06667	.26825	.967	-.7184	.5850
	Prettest of Control group	.20000	.26825	.738	-.4517	.8517
Prettest of opinion exchangetasks group	Prettest of information gap group	.06667	.26825	.967	-.5850	.7184
	Prettest of Control group	.26667	.26825	.585	-.3850	.9184
Prettest of Control group	Prettest of information gap group	-.20000	.26825	.738	-.8517	.4517
	Prettest of opinion exchange tasks group	-.26667	.26825	.585	-.9184	.3850

As the above table shows, there is no significant difference among the groups. The same statistical procedures were used for the post-test results. The results of statisticalanalysis of the post-test for the experimental and the control group are presented in Table 5.

Table 5

Descriptive Statistics for the Participants' Performance on the Post-test

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
Posttest of information gap group	30	2.4667	.74322	.19190	2.0551	2.8783	1.00	4.00
Posttest of opinion exchange tasks group	30	3.8000	.41404	.10690	3.5707	4.0293	3.00	4.00
Posttest of Control group	30	1.6000	.73679	.19024	1.1920	2.0080	.00	3.00
Total	90	2.6222	1.11373	.16603	2.2876	2.9568	.00	4.00

The mean score of the first experimental group rose from 1.33 on the pre-test to 2.46 on the post-test, which shows improvement. The mean score of the second experimental group, increased from 1.40 on the pre-test to 3.80 on the post-test, which indicates an improvement more than what is observed in the first experimental group. The control group's mean score was also improved a bit. In order to find out whether there are statistically significant differences on the learners' performance in three groups, the post-test scores were submitted to a one-way ANOVA analysis with between-group factor. The results are shown in Table 6.

Table 6

One-way ANOVA for Comparing the Performance of Groups (Post-Test)

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	36.844	2	18.422	43.632	.000
Within Groups	17.733	42	.422		
Total	54.578	44			

As table 6 shows, the results (Sig=000) illustrated that the difference betweenthe performance of three groups is statistically significant. In other words, information adopinion gap activities had a supportive role in communicative learning. Toilluminate where the significant differences fall among the groups, post hoc test(with an alpha level of .05) was conducted. The results are shown in Table 7.

Table 7

Post Hoc and Multiple Comparison of Three Groups

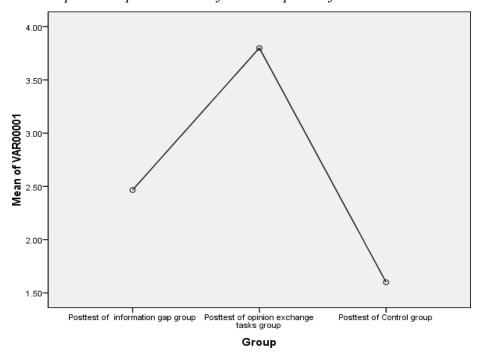| (I) Group	(J) Group	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval Lower Bound	Upper Bound
Post-test of information gapgroup	Post-test of opinion exchange tasks group	-1.33333*	.23727	.000	-1.9098	-.7569
	Post-test of Control group	.86667*	.23727	.002	.2902	1.4431
Post-test of opinion exchange tasks group	Post-test of information gap group	-1.33333*	.23727	.000	.7569	1.9098
	Post-test of	2.20000*	.23727	.000	1.6236	2.7764

		Control group					
Post-test of Control group	Post-test of information gap group	-.86667[*]	.23727	.002	-1.4431	-.2902
	Post-test of opinion exchange tasks group	-2.20000[*]	.23727	.000	-2.7764	-1.6236

The results revealed that information-gap and opinion-exchange outperformed the control group. It can be concluded that information-gap and opinion exchange activities had a significant and meaningful effect on participants' learning, but opinion-exchange task was more effective. Figure 2 shows the graphical comparison of the two groups on the post-test.

Figure 2

The Graphical Representation of the Groups' Performance on the Post-test

Discussion

In this section, the obtained results mentioned above are discussed in order to find answers to the research questions.

Addressing the First Research Question

The results revealed that the participants' WTC was enhanced, e.i. information-gap tasks were useful in improving the learners' WTC in EFL context. This is in line with a number of studies which have come to the conclusion that information-gap tasks have positive effect on enhancing teaching and learning practices. Fallahi et al. (2015), for instance, found that information-gap tasks were very effective on EFL learners' reading comprehension ability. The finding is also in line with that of Fatemipour and Mohammadi (2014)who specifically investigated the impact of using information-gap activities on improving EFL elementary learners' willingness to communicate. They reached a significant effect of these tasks on enhancing EFL learners' WTC, too. This finding of the present study also lends support to those of Kang and Pica (2006) who claimed that information gap tasks were a crucial component, and Fulmer (2010) who observed that significant differences exist between those who use tasks for enhancing theirwillingness to communicate in daily life and those who do not.

Addressing the Second Research Question

Regarding the second research question, 'Do opinion-exchange tasks have any significant effect on Iranian EFL learners' WTC?', the obtained results proved to the positive and thus support the findings of previously-conducted studies. For example, Fallahi et al. (2015) revealed that the opinion-exchange task was very effective on reading ability for a few reasons. Marzban and Hashemi (2013) too concluded that opinion-exchange task boosts Iranian intermediate EFL learners speaking ability. They reported that students who used opinion-exchange tasks were more motivated and interested in speaking than the others.

Addressing the Third Research Question

The results revealed that opinion-exchange task was more effective than information-gap task. This can be due to the belief that this type of task requires the learners' desire to speak their minds about a topic.So, the learners find the need to say something and this is where this need demands them to be more willing to communicate. The same finding was obtained in Fallahi et al.'s (2015) study which was carried out in an Iranian context to see if the impact of information-gap and opinion- exchange tasks have any effects on reading comprehension improvement of

Iranian EFL learner. These researchers found that opinion-exchange task was more effective than the information-gap task for a few reasons: They claimed that they observed real personal involvement, with an accompanying increase in confidence and fluency. Further, students talked about their opinions and preferences while at the same time wanted to convince their partners that their idea was the best idea.

Conclusion

The present study was an attempt to investigate the effects of two tasks proposed by TBLT scholars on Iranian EFL learners' willingness to communicate (WTC). According to the results, it was observed that the tasks under the study had positive impacts on the learners' willingness to communicate, and that through information-gap tasks, the participants demonstrated more willingness to communicate than members of the control group. Regarding the second type of task, i.e. opinion-exchange task, the participants showed far more willingness to communicate in the sense that they were too enthusiastic to speak their minds. Moreover, their attempt to advance in proposing their opinions was noteworthy. It was, in fact, revealed that they indicated far more willingness for communication than the participants of the other experimental group (information-gap group).

The findings of this study can be taken into account by Iranian EFL teachers in taking tasks to their classrooms. They can be a help in making links between TBLT to promote EFL learners' willingness to communicate. Also, opinion-exchange tasks can be used by EFL teachers in classrooms with higher levels of proficiency, since these students are more proficient and have no problem speaking in English. In addition, the findings can be used by textbook designers and syllabus practitioners in incorporating tasks which better render communication. In so doing, they can use information-gap and opinion-exchange tasks since this study, along with previous studies, proved that they can result in fair amount of willingness to communicate. Moreover, teachers can use the present findings in their classrooms to have a more dynamic environment which submit better results.

References

Akbarzadeh, M., & Narafshan, M. H. (2016). A Study on the Relationship between EFL Learners' anxiety and Willingness to Communicate in Language Classes. *International Journal of*

Language Learning and Applied Linguistics World, 11(2), 73-85.

Breen, M. P., Ed. (2001). *Learner Contributions to Language Learning: New Directions in Research Applied Linguistics and Language Study*. Routledge, England.

Ellis, R. (2009). Task-based language teaching: Sorting out the misunderstandings. *International Journal of Applied Linguistics*, *19*(3), 221-246.

Ellis, R. (2009). The differential effects of three types of task planning on the fluency, complexity, and accuracy in L2 oral production. *Applied Linguistics*, amp 042.

Fallahi, S., Malayeri, F. A., & Bayat, A. (2015). The Effect of Information-gap vs. Opinion- gap Tasks on Iranian EFL Learners' Reading Comprehension.

Fatemipour, H., & Nourmohammadi, A. (2014). The impact of using information-gap activities on improving EFL elementary learners' willingness to communicate. *Social Research Quarterly*, *6*(4), 109-123.

Jahanshahi, E. (2013). *The Effect of Group Work on Learners' Willingness to Communicate in EFL* (Doctoral dissertation, Shahid Rajaee Teacher Training University).

MacIntyre, P. D., Baker, S. C., Clément, R., & Conrod, S. (2001). Willingness to communicate, social support, and language-learning orientations of immersion students. *Studies in second language acquisition*, *23*(03), 369-388.

MacIntyre, P. D., Dörnyei, Z., Clément, R., & Noels, K. A. (1998). Conceptualizing willingness to communicate in a L2: A situational model of L2 confidence and affiliation. *The Modern Language Journal*, *82*(4), 545-562.

Maftoon, P., & Sarem, S. N. (2013). Gender and Willingness to Communicate.*Iranian Journal of Language Issues (IJLI)*, *1*(1).

Marzban, A., & Hashemi, M. (2013). The Impact of Opinion-gap Tasks on the Speaking of Iranian Intermediate EFL Learners. *Procedia-Social and Behavioral Sciences*, *70*, 943-948.

Montasseri, Z., & Razmjoo, S. A. (2015). The Effect of Using Competitive and Cooperative Teaching on the WTC of Iranian EFL Learners. *International Journal of Language and Applied Linguistics*, *1*(3), 54-61.

Murad, T. M., & Smadi, O. (2009). *The effect of task-based language teaching on developing*

speaking skills among the Palestinian secondary EFL students in Israel and their attitudes towards English. Department of Curriculum andInstruction Faculty of Education Yarmouk University.

Nunan, D. (1991). Communicative tasks and the language curriculum. *TESOL quarterly*, *22*(5), 279-295.

Nunan, D. (2005). An introduction to Task-Based Language Teaching. *The Asian EFLJournal Quarterly*, *7*(1), 25-28.

Richards, J. C., & Renandya, W. A. (2002). *Methology in language teaching: An anthology of current practice*. Cambridge university press.

Song, M. and S. Zhang, (2008). EFM: A Model for Educational Game Design, in Lecture Notes in Computer Science, Springer US. p. 509-517.

Trevens, P. (1992). Applied linguistics: An overview. *Introduction to applied linguistics*, 13-31.

Zarrinabadi, N. (2014). Communicating in a second language: Investigating the effect of teacher on learners' willingness to communicate. *System*, *42*, 288-295.

Letter to the Editor

Social Isolation and Disconnectedness of Translators: An Overlooked Occupational Concern

Mehrdad Vasheghani Farahani [1*] & Nematoolah Shamoosi [2]

[1]Department of Applied Linguistics and Translation Studies, Leipzig University; Germany

[2]Sabzevar University of Medical Sciences, Sabzevar, Iran

Dear Editor in Chief

This letter comes to you in the hope of highlighting a concern about the affective side of translators' occupation. In fact, by getting focused on sensitive texts such as law and medicine (Drugan, 2017) and spending long hours in isolation for deeper concentration, translators get engaged in challenging intellectual work, which is an inherent characteristic of translating from one language and culture to another (Drugan, 2017). Accordingly, they will miss opportunities to be in nature, to join with friends and relatives, to get engaged in social events in addition to losing chances of intimacy (De Jong Gierveld & Van Tilburg, 2010). As a compensatory strategy for decreased chances of social life, online friendships and relations develop for translators, as in most other contemporary professions. In following lines, we will briefly see to the strengths and disadvantages, and suggest ways to draw further attention to this affective concern for translators.

On the one hand, engaging with the audience in written discourse (Hyland, 2001 & 2005) provides a basic taste of sociability for an author of a text, but unregulated occupations such as translation naturally urge translators to spend hours in isolation and at a distance from others to

---

[1]Corresponding Author's Email:
mehrdadfarahani1365@gmail.com

concentrate on accuracy of the equivalence, smooth written text, and polished production for commissioners and readers. In fact, this might even lead them to develop a sense of professionalism by experience, and cross-cultural orientation by exposure to diverse disciplinary and interdisciplinary texts. In other words, translators involved in precise intellectual trans-creation of the original texts are often seen to value lifelong learning and self-development (Katan, 2009), while being faithful to the text and sacrificing their lives to transferring knowledge from a source language to the readers of the target language.

Relevant occupational theories stress the need for social connectedness as a deeply ingrained human characteristic (Cacioppo & Hawkley, 2009). As a matter of fact, participation in social activities and social network ties are associated with better mental health (Cornwell & Laumann, 2015, Vasheghani Farahani & Shomoossi, 2021). Despite sporadic social activities, occupational concentrations leading to translators' loneliness or feeling isolated are more likely to cause persistent fatigue symptoms such as depression (Chen & Feeley, 2014) which may manifest both physically and cognitively (Ocon, 2013).

Despite its being considered as probably the "second oldest profession" (Baer & Koby, 2003: viii), translators even admit being deprived of a societal recognition status (Katan, 2009). While Sela-Sheffey (2008) laments the lack of research or findings regarding their social status, she does suggest that "all evidence shows that [translators] are usually regarded as minor, auxiliary manpower" (p.2). Despite developing their own perspectives by education and career experience, such paradoxical effects may further lead translators to perceived identity loss.

In short, we wrote this letter to highlight a concern over the mental health of translators during their life course. Isolation from social life, worries about one's societal status, concentrating on translation even after the office hours, sedentary lifestyle due to overwork, and lifelong commitment to translation are felt to aggravate factors in ill-health both physically and cognitively. Therefore, appropriate regulation and enhanced professional advantages such as payment and insurance to compensate for perceived social disconnectedness can be topics for further research in this area of inquiry. Future studies are necessary to formally determine relevant variables and aspects. Sustainable strategies may be empirically developed to address affective problems, in particular.

References

Cacioppo, J. T., & Hawkley, L. C. (2009). Perceived social isolation and cognition. *Trends in cognitive sciences*, *13*(10), 447–454. https://doi.org/10.1016/j.tics.2009.06.005

Chen, Y., & Feeley, T.H. (2014). Social support, social strain, loneliness, and well-being among older adults: an analysis of the Health and Retirement Study. *The Journal of Social and Personal Relationships*, *31*(2), 141-161. doi: 10.1177/0265407513488728.

Cornwell, B., & Laumann, E. O. (2015). The health benefits of network growth: new evidence from a national survey of older adults. *Social Science and Medicine* (1982), 125, 94–106. doi: 10.1016/j.socscimed. 2013.09.011

De Jong Gierveld, J., & Van Tilburg, T. (2010). The De Jong Gierveld short scales for emotional and social loneliness: tested on data from 7 countries in the UN generations and gender surveys. *European journal of ageing*, *7*(2), 121–130. https://doi.org/10.1007/s10433-010-0144-6

Drugan, J. (2017). Ethics and social responsibility in practice: interpreters and translators engaging with and beyond the professions. *The Translator*, 23(2): 126-142Special issue: Translation, Ethics and Social Responsibility. Doi: 10.1080/13556509.2017.1281204

Hyland, K. (2001). Bringing in the reader: addressee features in academic writing. *Written Communication, 18*(4): 549-74. Doi: 10.1177/0741088301018004005

Hyland, K. (2005). Representing readers in writing: student and expert practices. *Linguistics and Education*, 16(4), 363-377. Doi: 10.1016/j.linged.2006.05.002

Katan, D. (2009). Occupation or profession: A survey of the translators' world. translation and interpreting studies: *The Journal of the American Translation and Interpreting Studies Association*, 4, 187-209. DOI: https://doi.org/10.1075/tis.4.2.04kat

Ocon A. J. (2013). Caught in the thickness of brain fog: exploring the cognitive symptoms of chronic fatigue syndrome. *Frontiers in physiology*, 4, 63. https:// doi.org /10.3389 /fphys. 2013.00063.

Vasheghani Farahani, M., & Shamoosi, N. (2021). Teleworking translators during the COVID-19 pandemic: Social and personal experiences from Iran. *New Voices in Translation Studies*, 24, 117-126.