



## Iranian Teachers' Views on Educational Philosophies of Dewey, Rousseau and Aristotle

Mahmood Reza Atai<sup>1</sup>

Farhad Mazlum

*Teacher Training University, Tehran*

### Abstract

The present study aimed at investigating Iranian teachers' views about and reactions to the educational philosophies of Dewey, Rousseau and Aristotle. Based on the most important educational philosophies and thoughts of these thinkers, a likert-scale questionnaire was developed. Four open-ended items were also added to let the answerers express their personal philosophies freely. One hundred and twelve teachers participated in this study. Statistical analyses of data (i.e., factor analysis and repeated measure ANOVA) revealed that the respondents had reacted differently to Aristotelian education vis-à-vis Dewey's and Rousseau's educational beliefs. The Aristotelian 'spreading morality is the main purpose of education' was significantly rejected. Although there was a statistical difference in data concerning Dewey's vs. Rousseau's educational viewpoints, it was not significant. To shed some light on the findings and to explain the participants' inclination toward Dewey's and Rousseau's educational understandings and their general disapproval of Aristotelian education, a close scrutiny of some specific items of the questionnaire and open-ended items was followed. The participants' educational views and philosophies are discussed and compared with the educational views of these thinkers. The results are revealing and informative since Western educational philosophies are studied in an Eastern Islamic context.

**Keywords:** Educational philosophy, Dewey, Rousseau, Aristotle, Iranian teachers' philosophy

### 1. Introduction

Crews and Glascott (1998) maintain that "your philosophy will lead you" (p. 234). That is why scholars such as Ryan (2007) argue that it is important to study teachers' educational philosophy since it not only guides their actions but it also "motivates revisions and grounds teachers throughout their careers" (p. 44). In a similar vein, Petress (2003) holds that a teacher's educational philosophy "constitutes a moral and social compass, behavioural, attitudinal, and value guide-posts, essential personal and professional evaluation" (p. 1). All this explains the significance of investigating and understanding what is nowadays called teachers' *philosophical identity*.

---

<sup>1</sup> Email: mreatai@yahoo.com



Different teachers may teach in a certain manner for many unexamined and inaccessible reasons. Levin and Nolan (2004), however, suggest that “the single most important factor in determining the learning environment is teacher behavior. Intentionally or unintentionally, teachers’ verbal and nonverbal behaviours influence student behaviours” (p. 1). Arguably, all teacher behaviour is connected to his or her philosophy, beliefs, values and motivations which “... become visible during teaching because teaching demands action, reflective thought and revision as success is pursued and challenges are expected” (Ryan, 2007, p. 40). Teachers as key facilitators reveal the curricula content to their students by their behaviours and actions. Briefly then, teacher behavior is connected to his educational philosophy and, at the same time, it affects student behaviour.

Each teacher’s conception of philosophy is informed by beliefs about human existence, educational product and process. According to Ryan (2007),

These beliefs nourish our positions as we may be called on to speak about the issues relating to the meaning of life, our universe, the limits of knowledge and truth while teaching, coaching, counseling or guiding students. The need to respond to these issues requires self-knowledge, skills and sensitivity (p. 41).

Petress (2003) maintains that philosophy “infuses our moral and social compass, behavioural, attitudinal, and value guide-posts, essential personal and professional prescriptions, and a consistent but alterable assessment means for professional evaluation” (p. 1). Actually, educators’ philosophy impacts and affects perceptions, beliefs, understanding and values to the points where all decisions can be traced back to their educational philosophy.

Historically, some educational philosophers seem to be more outstanding due to the influences their ideas have had. Socrates, Plato, Aristotle, Rousseau, Pestalozzi, Herbart, Froebel, Dewey, Priere are among the most well-known thinkers who brought about shifts, sometimes big ones, in both philosophical orientations and practical issues. Some of the changes they made are still highly appreciated, some are critically questioned, and some totally refuted.

This study was an effort to investigate and unearth Iranian teachers’ views on and reactions to educational philosophies of Dewey, Rousseau, and Aristotle. Although space does not permit an extensive elaboration of their educational philosophies, a very short review is provided below.

### *1.1. Aristotle*

Aristotelian educational ideas are still relevant for today’s educators (Noddings, 1998). For Aristotle, man is not good by nature. Therefore, the purpose of education is to produce good and virtuous



men. Generally speaking, Aristotle's name is associated with *moral education* and *reasoning*. The most important educational beliefs of Aristotle that were taken into consideration in this study include the following: (1) The fundamental aim of education is to spread morality; (2) Initially, children cannot reason; (3) During the early years, the focus should be on children's right conduct; (4) Community needs and welfare sometimes override individual rights and children's education must be adjusted to the needs of society; (5) People should be educated or trained for their appropriate places in life; (6) People cannot guarantee their moral behavior because (7) the environmental factors and circumstances make people behave unprecedently.

### *1.2. Jean-Jacques Rousseau*

In educational philosophy, Rousseau is referred to as the *philosopher of freedom* (Noddings, 1998) because he seemed to "...extol the natural (or primitive) state of human beings over the civilized one, and in nature, human beings ... are free of the pressures and corruptions of the political state" (Noddings, p. 14). For Rousseau, children are born good and do not need the rigid moral education that Aristotle proposes. The following educational views and beliefs of Rousseau were taken into account in this study: (1) Men are born free and good; (2) Children need no rigid moral education; they are already good; (3) Nature should be the main source of educational inspirations; (4) Children's motivation, direct action, senses and feelings are primary in educating them; (5) Children's interest is the key determinant of type of education; (6) Timing in education is important (i.e., children are ready at certain times to learn certain things); (7) The entire education of women must be relative to men (i.e., to please them, to be useful to them, to be loved and honored by them); (8) The right direction in education is to move children toward happiness and freedom; and, (9) Through exerting more control and by manipulating environmental factors, better results would be obtained.

### *1.3. John Dewey*

Dewey is known as a *naturalistic* philosopher who believed in what he called the *method of science* and advocated its use in every sphere of human activity. His educational views have often been accompanied with controversies (Noddings, 1998). Among his educational views, the following were considered in this study: (1) Children are not born good; rather they have the potential of becoming good or bad; (2) Motivation and learning-by-doing are the main factors in learning process; (3) Students must be involved in the establishment of objectives for their own learning; (4) Education must take students' prior experiences into account and be in concert with them; (5) Different students need different kinds of education; (6) Teachers should not impose beliefs and ideas; (7) Making students learn different subjects (e.g., history, reading, science, geography...)



simultaneously would interfere with their *natural growth*; (8) Education does not mean preparing and fitting children for future life; education is life itself.

## 2. Method

This study aimed at investigating the Iranian teachers' reactions to the educational philosophies and thoughts of Aristotle, Rousseau, and Dewey. In so doing, some of the most important educational philosophies of these philosophers were selected, put into simple statements, translated into Farsi, reviewed and revised several times after critical and constructive comments of the departments' professors, and finally, were put into a 28-item questionnaire.

### 2.1. Instrument

A 28-item questionnaire was developed on the basis of educational thoughts and philosophical orientations of Aristotle, Rousseau, and Dewey. Nine items contained Rousseau's philosophical ideas, twelve items described some of Dewey's educational orientations, and the other seven items included some educational beliefs of Aristotle. The likert-scale questionnaire required the respondents to choose from Strongly Agree, Agree, Neutral, to Disagree. The fifth column; i.e. comments or ideas, provided the respondents with the opportunity to express their further comments or explanations about some possibly controversial or debatable notions freely. Furthermore, and to gain more revealing explanations, four open-ended statements were added to the questionnaire that asked the participants to elaborate on the following: (1) *The most effective way(s) of educating Iranian students...*; (2) *The major aim of education is...* ; (3) *To play an efficient role in students learning, a teacher should have the following features...* (4) *Learning would be more efficient if a student has the following characteristics...*

### 2.2. Participants

Two hundred questionnaires were distributed to public school teachers in Maragheh and Tehran. They were male and female teachers at different levels and of varying majors. Most of the respondents (i.e., 94%) held a BA or a BS. The participants taught in both rural and urban schools. Only 139 questionnaires were given back out of which 112 questionnaires were selected for data analyses.



### 3. Data Analysis

Factor analysis revealed three loadings of which factor three can be easily and systematically followed among the items. All the items rotated under this category belong to the items pertaining to the general inclusive category of '*Moral Education*'. The second factor -though not as neatly rotated as the first one- consists of items reflecting some principles of '*Learner-centered Education*': the importance of motivation, feelings, senses, prior learning experience, hands-on-activities and fun in learning process. No label can be assigned to the first factor due to the mismatch of varying items under this loading.

Bearing in mind the main purpose of the study; i.e. Iranian teachers' views about the educational thoughts and philosophies of Aristotle, Rousseau, and Dewey, Repeated Measure ANOVA was run to see if participants of this study reacted differently to these philosophers' educational views.

Eyeballing the descriptive statistics shows the mean differences of each philosopher from the other suggesting that Dewey and Rousseau have been reacted to almost similarly. Teachers' reactions to Aristotle's views, however, have resulted in a markedly lower mean. Table 1 below shows the descriptive statistics of the data.

**Table 1.**

Means of the respondents' views on educational views of Rousseau, Dewey, and Aristotle

	Mean	Std. Deviation	N
Rousseau	26.1607	2.74290	112
Dewey	26.7232	3.99822	112
Aristotle	20.0804	3.04364	112

Further statistical analyses revealed that the mean differences were statistically significant at the .05 level but the difference lied with Rousseau vs. Aristotle and Dewey vs. Aristotle. Put in simple terms, the mean difference was the result of the participants' meaningful different reaction to Aristotle's standpoints. They, nevertheless, seem to have reacted almost similarly to the educational philosophies of Dewey and Rousseau. Though Dewey's mean was higher than that of Rousseau, it was not statistically significant. Tables 2 and 3 show the results of the tests of within-subject contrasts and pairwise comparisons respectively.



**Table 2.**

Tests of within-subjects contrasts

Source	philosop	Type III Sum of Squares	df	Mean Square	F	Sig.
philosop	Linear	2070.362	1	2070.362	349.182	.000
	Quadratic	969.121	1	969.121	136.911	.000
Error(philosop)	Linear	658.138	111	5.929		
	Quadratic	785.713	111	7.078		

**Table 3.**

Pairwise comparisons

Philosoph (J) philosoph	Mean difference (I-J)	Std. Error	Sig.*	95% Confidence Interval for Difference <sup>a</sup>	
				Lower Bound	Upper Bound
2	-.563	.325	.086	-1.206	.081
3	6.080*	.325	.000	5.436	6.725
1	.563	.325	.086	-.081	1.206
3	6.643*	.370	.000	5.909	7.377
1	-6.080*	.325	.000	-6.725	-5.436
2	-6.643*	.370	.000	-7.377	-5.909

Based on estimated marginal means

\*. The mean difference is significant at the .05 level.

a. Adjusted for multiple comparisons: Least Significant Difference (equivalent to no adjustments).

## 4. Results

The findings of the study tend to suggest that the respondents have reacted differently to the educational thoughts and philosophies of Aristotle compared to those of Dewey and Rousseau. More specifically, it was found that the difference in reaction to Aristotle's views vis-à-vis Dewey and Rousseau's ideas was statistically significant.

Based on the findings of the study several points are worth noting. First and foremost is the participants' reaction to Aristotle's educational views and thoughts. Most of the items reflecting Aristotle's views can be put under the general label *moral education*. For him, the primary purpose



of education is to spread morality and produce virtuous and pious men. For teachers in this study, nevertheless, *spreading morality* is not the only and the main function of an educational system. Scrutinizing the answerers' feedback on the items pertaining to Aristotle's ideas makes it evident that the lowest average, hence the statistically significant difference, is due to the respondents' choice of 'Neutral' or 'Disagree' in large numbers. Interestingly, items reflecting Aristotle's viewpoints have received the most 'Neutrals' yielding to the lowest mean. Does it mean that the participants have been conservative in answering 'morality' items, despite the fact they knew anonymity was taken into account? In the fifth column of the questionnaire, ideas and comments, some teachers stated that morality is not/ should not be the major concern of an educational system, since moral values make up only a part of the whole.

Secondly, the participants seem to have approached Dewey and Rousseau's ideas and philosophies more or less similarly. In other words, these philosophers' ideas have not been distinguishing enough to lead to a well-observed and significant separation or reaction. These philosophers share some of the items concerning the importance of motivation, learning-by-doing or experiential learning and they differ in few items; e.g. people are born 'free' (Rousseau) vs. people are not born 'free'; they have the potential for being both good and bad (Dewey). A close look at the remaining items -which are more than the items reflecting points of departure-, sheds enough light on the findings. These items have largely been appreciated or shunned generally unanimously. That is the respondents have generally been in agreement or disagreement when reacting to these remaining items. What matters most is what these items are about. In addition to the importance of motivation, learning-by-doing or experiential learning for the respondents, the following principles are important for the participants.

- Different students should have different educations.
- Meaningful learning experiences are more efficient.
- Students' feelings play an important role in education.
- Environmental factors can be manipulated for the betterment of education.

The close mean indices for these two thinkers and the insignificant difference, however, does not mean that the respondents have agreed (or disagreed) with Dewey and Rousseau completely similarly and on all items. There are items for each philosopher that have been reacted negatively. Dewey, for example, holds that making students learn different subject mattes simultaneously would interfere with their 'natural growth'. Only 38% of the answerers have shown their agreements with this very item, putting 'Strongly agree' and 'Agree' together. 38% of them have disagreed with Dewey. Discrepancies like this between Dewey and the respondents -and between Rousseau and



the respondents- have been few leading to the overall unanimity and harmony of the participants' reactions.

Though the respondents have manifested a weaker propensity towards Aristotelian moral education, they have strongly disagreed with Rousseau on his 'Children do not need moral education' item. 83% of the participants have reacted negatively to this item which indicates that they do not believe in dispensing with moral education totally. As said before, some respondents have explicitly stated that 'spreading morality' is not the only game in town. Their reaction towards Rousseau's extreme stance might suggest that although they maintain that morality is not be the main purpose of education, it should not be neglected either.

#### *4.1. Items calling for attention*

Some specific items have been reacted to in such ways that a one-by-one in-depth analysis of them might prove revealing, informative and at the same time intriguing. Though Rousseau's views have been generally agreed with, the item concerning with male-priority education has been disfavored strongly. Rousseau holds that girls' education should be in accordance with what boys need, expect them to fulfill... A large group of the participants (i.e., more than 94%) have refuted such a sex-bias approach towards education.

The next two items- one reflecting Rousseau's idea about being 'free' and 'good' at birth and the other belonging to Dewey arguing that man is not born free and 'good' at birth, he has the potential for being both 'good and bad'- have been reacted to quite differently. Putting 'Strongly agree' and 'Agree' together, only 33% of the answerers are in line with Dewey's viewpoint whereas it is 95% for Rousseau's. About 43% of the answerers have opposed Dewey. The strong preponderance of Rousseau in this particular case might be attributable to the participants' common religious background- Islam.

Item 25 states that children are not capable of 'reasoning'- such sophisticated capability emerges in twenties. Disparity among the respondents can be easily detected if one compares 37% of disagreement with 49% of agreement- adding 'Strongly agree' and 'Agree' together.

#### *4.2. Exploration of open-ended statements*

Four open-ended statements were added to the questionnaire. Not all 112 respondents commented on these statements. The recurring themes in teachers' answers were classified and given percentages.



#### *4.2.1. The most effective way(s) of educating Iranian students...*

The most frequent theme (i.e., 43%) in participants' answers is "learning-by-doing". This principle is the core of what is generally known as experiential learning; itself an important mechanism of Progressivism. This principle is generally associated with Deweyian educational philosophy in which engaging students in hand-on activities is highlighted.

In the second place stand "affective" considerations. This group of participants (about 23%) argued that factors pertaining to students' affective domain play a key role in their learning and, therefore, the most effective way to provide learning opportunities is to prepare students affectively and emotionally. As went earlier, this argument is close to Rousseau's educational views.

Less frequent answers include team work, discussion (in opposition to lecturing), Islamic education, and others.

#### *4.2.2. The major aim of education is...*

For most of the participants of this study, an educational system should "prepare children for future life". This is the most recurring theme in data and an indication of the fact that education, for the respondents, is a means to an already defined end. This is in sharp contrast with Dewey's educational views. For Dewey, education is synonymous with growth and the aim of education is more education. Secondly, the respondents do not consider "developing morality and religion" as the major aim of education. This further confirms their responses to the items pertaining to Aristotelian educational philosophy.

#### *4.2.3. To play an efficient role in students learning, a teacher should have the following features...*

The answers to this item call for further studies so that more information is obtained about teachers' self-perception. Although teachers in this study had expressed their agreement with Rousseau's views on the importance of motivation, students' direct action, their interests and happiness, they see no role for themselves in motivating students or engaging them in actions or moving them on happiness track. Instead, they emphasize professional qualifications and character. More specifically, the two dominant themes in data reveal that about 34% of the participants regard teachers' field-specific knowledge and expertise as the most significant features of an efficient teacher in educational system. 30% of the answers deal with a teacher' character: he should be patient, nice, and kind.



Similarly, it can be argued that despite their agreement with some Deweyian tenets in the questionnaire (e.g., motivation and learning-by-doing are the main factors in learning process; and, teachers should not impose beliefs and ideas), teachers in this study seem to believe that motivating students and helping students learn by doing should be undertaken by others in the educational system.

#### *4.2.4. Learning would be more efficient if a student has the following characteristics...*

Teachers' perception of the role of students is also important and is considered one of the main components of teachers' philosophical identity. This item was to examine such perceptions. Politeness comes first (28%), followed by hard work and perseverance (16%), punctuality and discipline (13%), and being purposeful (13%). Interestingly, quite few teachers mention characteristics such as "critical and questioning" (2%) or "happy" (4%).

The findings seem to complement teachers' views and understandings of themselves discussed above. They laid a great emphasis on their professional qualifications and put it at the pinnacle. At the same time, they maintain that the efficiency of the learning process will increase if students are polite, punctual, perseverant, etc. This suggests that the old orthodoxy of teachers as "knowledge transmitters" and students as "knowledge receivers" still prevails among some teachers. In other words, teachers attribute a passive image to their students in the learning process: they are there to absorb the knowledge.

## **5. Discussion**

A teacher's philosophy is a thing of constant change because there is never an absolute when it comes to teaching students. Every classroom is different and each student has different needs. Beliefs and values regarding education and teaching affect a teacher's classroom practices. His beliefs and values, at the same time, are affected by outside variables and factors. This ongoing dynamism in varying contexts would work out more efficiently compared to static attitudes of some teachers who would either *impose* their educational thoughts or would *admit* to the learning setting naively. As life-long learners, teachers need to cultivate, refine and enrich their educational philosophies. The spin-off would naturally be more effective pedagogical practices.

One source of the ongoing and developing educational philosophy could be the educationalists' and philosophers' ideas and thoughts. Some philosophers and thinkers have affected our theories and consequently our practices greatly. Teachers might be conscious of the ideas and views they pursue or they might go after them without knowing whose thoughts they are. This study focused on



some of the important educational views of Dewey, Rousseau and Aristotle. The results tended to suggest that the Aristotelian education has been reacted to differently in comparison with Dewey's and Rousseau's views. It was found that there was an inclination toward Dewey's and Rousseau's educational understandings and a general disapproval of Aristotelian education. It was also maintained that some of Dewey's and Rousseau's were disfavored by the participants.

Answers to open-ended statements revealed that for the participant teachers, education is a means to an end: it should fit and prepare students for future life. To fulfill such a goal, teachers highlight the significance of teachers' professional qualifications and their compassionate, kind, and nice character. They also maintain that students should be polite, punctual, perseverant, etc so that the learning process gets facilitated. Despite their agreement with the significance of "learning-by-doing", "motivation", "happiness", etc in learning process, teachers in this study seem to consider other officials responsible for taking care of such issues.

This study raised a more fundamental question: Iranian teachers' philosophy of education. In other words, there are no studies, to the authors' best knowledge, investigating Iranian teachers' educational philosophies. As discussed before, educational philosophy acts like a compass; it guides teacher behaviour and affects curriculum implementation. Therefore, it is indeed important to know and study Iranian teachers' philosophical identity. The following two excerpts are intentionally selected to demonstrate how much teachers differ in their perceptions and views toward educating children.

*A child is inundated with energy. A child tells something about Allah's beauty, magnificence and power. To feel this, the teacher needs to beautify his mentalities and smarten his educational landscapes. He should be generous with his time and patience. Life cannot be taught to the child since life is an art- the kind of art that cannot be transmitted. Necessary means should be provided and he would learn it, make it or discover it himself. Education is helping with that making... that discovering.*

(Questionnaire No. 21)

*We should go back to the educational system we had before the revolution. It works well with Iranian students. I mean, that system suits and fits them. Also, to maintain discipline, corporal punishment should be allowed.*

(Questionnaire No. 75)



## References

- Beck, C. (1993). Postmodernism, pedagogy, and philosophy of education. *Philosophy of Education Yearbook* [electronic version]. Retrieved January 26, 2008, from <http://www.ed.uiuc.edu/EPS/PES-Yearbook/93-docs/>
- Crews, N., & Glascott, K. P. (1998). A teaching philosophy: Rhetoric or reality? *Childhood Education*, 74, 232-234.
- Dewey, J. (1916). *Democracy and education*. New York, NY: Macmillan
- Levin, J., & Nolan, J. F. (2004). *Principles of classroom management: A professional decision-making model*. (4th ed. pp. 234-250). New York: Allyn & Bacon
- Noddings, N. (1998). *Philosophy of education*. Westview Press.
- Petress, K. C. (2003). An educational philosophy guides the pedagogical process. *College Student Journal*, 37, 128-135.
- Phillips, D. C. (1998). John Dewey's the child and the curriculum: A century later. *The Elementary School Journal*, 98, 403-415.
- Ryan, T. (2007). The identification of pre-service teachers' philosophical orientation. *New Zealand Journal of Teachers' Work*, 4, 39-46.

## Bio Data

**Mahmood Reza Atai** is associate professor of applied linguistics at Kharazmi (Tarbiat Moallem) University, Tehran, Iran. He is the editor of *Iranian Journal of Applied Linguistics* and head of ELT & ESP Department at SAMT Organization. His research basically addresses academic literacy education and EAP issues. His recent papers appear in *System*, *Computer Assisted Language Learning*, *RELC Journal*, *Innovation in Language Learning and Teaching*, *Curriculum Journal*, *Asia TEFL*, and *Asian ESP Journal*. Also, he has published some EAP textbooks for Iranian university students.

**Farhad Mazlum** is a faculty member of Maragheh University. He has just finished his PhD program in Tarbiat Moallem University in Tehran. For his dissertation, he has been working on ELT policy, planning and practice in Iran. He serves on the editorial board of Asian EFL Journal.



## North-American and Iranian EFL Lecturers' Use of Discourse Markers: The Cross-cultural Aspects

**Zohre Eslami-Rasekh**

*Texas A&M University*

**Abbass Eslami-Rasekh**

*University of Isfahan*

**Shahla Simin<sup>1</sup>**

*University of Isfahan*

### **Abstract**

Lecturing is considered as an inevitable part of teaching practice in higher education. In non-native universities, lecturers need to be trained explicitly in how to use discourse markers (DMs) to make their speech comprehensible. To our knowledge, few research works focus on the pedagogical aspects of the use of these devices to make the unity and coherence of the speech explicit. This paper aims at comparing the English and Iranian EFL lectures to identify the frequency of use and various types of discourse markers by means of a corpus linguistics analysis. For this purpose, a corpus of ten spoken lectures, half (5 North-American English lectures, NAC) taken from MICASE (Michigan Corpus of Academic Spoken English) and the other half (5 Iranian EFL lectures, IC) recorded at one of the Universities in Iran. For analysis, the classification model proposed by Bellés-Fortuño (2006) was employed. The independent variable for the analysis was the language background of the lecturers and the dependent variable was the frequency rate of each pre-established marker in the two sub-corpora. Results of this comparative analysis showed that DMs were used more in the IC than in the NAC. An overall view on individual categories represented that micro-markers have got the highest rate in contrast with macro-markers and operators. This indicated that lecturers in both sub-corpora tended to express logico-semantic relations in their lectures explicitly. As a result, it could be said that in the discourse of lectures there was a need to convey lexical and descriptive meaning by using micro-markers. Analyzing collocational DMs illustrated that micro-markers tended to cluster with other markers more than macro-markers or operators. The reason could be due to the fixed and less variable structure of micro-markers, in contrast with the unsteady and less permanent structure of macro-markers.

**Keywords:** Corpus linguistics, Discourse markers, Micro-markers, Macro-markers, Operators

---

<sup>1</sup> Email: shahlasimin@yahoo.com



## 1. Introduction

In the process of globalization of languages, English as a lingua franca plays a vital role in commerce, business and academic world. This globalization, forces higher education institutions to internationalize and it is concerned with extensive student and staff exchanges. The joint development and implementation of new educational programs and recently, a worldwide focus on student and staff recruitment necessitates the application of a meta-language to be used for scientific advancements (Wilkinson, 2004).

Due to internationalization of higher education, several research projects have been conducted on the effect and use of English within academic discourse. Specially, studies on DMs and various aspects of lectures have raised scholars' interest (Chaudron & Richards, 1986; Flowerdew, 1994; Jones, 1999; Khuwaileh, 1999; Kerans, 2001; Thompson, 1994). Regarding DMs, several studies have proved the effectiveness of discourse markers for the comprehension of connected discourse. For instance, Fraser (1999) investigated DMs from a grammatical-pragmatic point of view and Lorente (1996) offered the idea of cataloguing DMs as 'cohesive links'.

Concerning different aspects of lectures, Murphy & Candlin (1979) recognized markers of the rhetorical organization of lecture discourse. Thompson (1994) investigated the discourse structure of lectures introductions and Young (1990, 1994) suggested five different phases for lectures: content, discourse structuring, conclusion, evaluation and interaction. All these studies focused on structure and relations within discourse.

In academic settings, teachers are involved with different instructional materials, such as seminars, tutorials, videos, and writing assignments, among others; but lecture "remains the central instructional activity" (Flowerdew, 1994). Benson (1994) defines lecture as "the central ritual of the culture of learning". All over the world, lecturing is considered as an effective practice in higher education (Dunkel & Davy, 1989). Understanding and comprehending lecture content seems to be necessary for students' university success; furthermore, students should be able to not only read English publications, but participate and deliver papers at conferences, even do research and give lecture in academic contexts.

In Iran, lecturers are not trained in how to give lecture nor are there specific trainings in regard to the use and functions of DMs which are fundamental for lecture comprehension. This study is an effort to analyze the content of lectures as an academic genre within a corpus linguistic approach. The aim is analyzing some linguistic features (DMs) that are useful in the understanding and



retention of lectures. The objective is to describe the use and functions of DMs and the similarities and differences that exist. The following research questions have been formulated:

- a) Are there any differences in the use of DMs including micro-markers, macro-markers, and operators between North-American and Iranian EFL monologic lectures in the discourse of Humanities and Social Sciences?
- b) Are there any specific collocational combinations serving as marking devices in North-American and Iranian EFL monologic lectures?

## **2. Corpus Studies on DMs**

Very few researches have been done regarding the spoken lecture discourse and the use and function of DMs. However, recently, researchers became interested in academic genres (seminar, conference presentations, etc.) specially the genre of lecture. DMs are essential for the understanding of written texts and they play a significant role in creating a meaningful and coherent message in the communication process of oral discourse.

The difference between spoken and written discourse affect the use and function of DMs. Those DMs which are more frequent in written texts are less frequently used in spoken discourse (*on the contrary/ to sum up, etc.*). Speakers have access to a richer context in spoken discourse, i.e., “they have prosody and phonology as well as non- verbal communication or interaction with external physical objects” (Bellés-Fortuño, 2006, p. 112).

The studies presented here have used spoken or written corpora as the method of study and they have focused on the use of a single or various DMs. In addition, these researchers provided a wide and broad classification of DMs. For instance, Del Saz (2003, 2005) focused on the notion of reformulation and the lexical units that explicitly convey reformulation. She called them Discourse Markers of Reformulation (DMs of RF). She claimed that what has been called reformulator can be considered a DM.

She analyzed naturally occurring instances of language collected in the British National Corpus (BNC) and considered reformulators as DMs because according to Fraser (1999) they have connectivity and non-truth conditionality which are features of a DM. The DMs of RF is based on Fraser's canonical form for a DM SEQUENCE, namely, S1-DM+S2. In her investigation, Del Saz (2003) attempted to fit DMs of RF into Fraser' (1999) classification.



Swales and Malczewski (2001) worked on spoken academic discourse and emphasized on what they called "a cluster of features that constellate around discourse management across a wide range of university speech events" (2001, p. 146). They described speech events as 'activity types' (Levinson, 1979) in which language is used to get things done. They pointed out a difference between academic speech that is monologic (often lectures, conference presentations, etc) and dialogic talk (telephone workshops, interviews, etc). They examined MICASE (Michigan Corpus of Academic Spoken English) (Simpson, Briggs, Ovens, & Swales, 2002) and studied the use of footing changes they called New Episode Flags (NEFs).

They focused on the linguistic resources (or NEFs) which participants used in different university events to move from lecture format to discussion (or the reverse) or change the direction of the lecture or discussion. They analyzed linguistic resources such as group vocatives (folks, gang, friends, guys), directive or vocative verbs (say, listen, look) and exhortative or jussive imperative let, all these not with a large number of occurrences. They also analyzed more frequent NEFs in the MICASE such as *okay*, *so* and *now* (DMs).

Rendle-Short (2003) is another researcher who studied academic discourse and examined the use of the DM *so* in seminar talk within the computer science discipline. She illustrated two functions of *so* in various contexts. Semantically, *so* connects adjoining clauses to show causal relations. Pragmatically, *so* marks the potential speaker discourse transitions functioning as a topic-shifter. Rendle-Short (2003) explored these functions of *so* in a corpus consisted of six first video-taped and then transcribed computer science seminar talks. The results showed that the DM *so* plays different functions in seminar talk depending on its position in the talk. She also proved that monologic talk is not continuous; in contrast, it is divided into smaller parts or sections which have a finely organized and well-structured discourse pattern.

### **3. Method**

#### *3.1. Corpus selection*

In this contrastive study, a corpus of ten spoken lectures was gathered. Half of the corpus (5 North-American English lectures) was taken from MICASE (Michigan Corpus of Academic Spoken English) (Simpson, Briggs, Ovens, & Swales, 2002) available on the Internet. MICASE contains a collection of transcripts of academic speech events recorded at the University of Michigan at Ann Arbor. The corpus consists of approximately 1.8 million words transcribed from a variety of speech events that goes from February 1998 up to 2003. The other half (5 Iranian EFL lectures) was recorded at Najafabad Azad University, Esfahan, (Iran). These lectures were selected randomly.



The whole lectures were recorded with an Mp3 recorder. Lectures were recorded under lecturers' permission.

The researcher attended the classrooms from April 2009 as an observer while lectures were delivered and took notes of the events that may affect lecture delivery such as: classroom distribution (male/female), number of students, number of students' interruptions, lecturers' use of resources (*Power Point Slides, handouts, etc.*). After recording the lectures, they were transferred to computer and saved as sound files. Then transcribing process began. For each lecture, a transcript draft and a CD of the recorded lecture were prepared. The transcribing process was a very arduous and time-consuming task. Then, the final transcript version was ready for the analysis. After analyzing the results, the researcher asked another person (an MA student) to proofread the findings.

### *3.2. Corpus description*

All the lectures chosen for this study belong to the division of Humanities and Social Sciences. This selection was made purposefully. Recently, Social Sciences as an academic field of study gained importance because the need for communicative competence has been increased. The reason is that teachers should gain the necessary features for the market demand in all oral genres, whether professional or academic, as it is the case of lectures. Another reason is related to the notion of 'neutrality'. It has been found that the discourse of Social Sciences is more neutral than that of Technical Sciences or the field of Humanities.

Thus, all the lectures in this study belong to the Academic division of Social Sciences. However, since the scope of Social Sciences is extremely large and it contains a wide range of disciplines such as Anthropology, Communication, Education, History, Psychology, etc., the researcher merely mentioned the title of each lecture to narrow the scope of the field for a better analysis. In this way, the field of Social Sciences is regarded as a single entity; so that, the aim of this study was achieved.

Important information about the lectures has been categorized according to the traits found in MICASE. Attributes like title, primary discourse mode, speech event, number of words, as well as recording duration are presented. *Primary discourse mode* refers to "the predominant type of discourse characterizing the speech event" (MICASE, Simpson, Briggs, Ovens, & Swales, 2002). In this study, only monologic lectures were selected where one speaker monopolizes the floor, sometimes followed by question and answer periods.



In the MICASE corpus *speech events* are classified according to classroom events and non-class events. This corpus consists of classroom lectures. Based on the number of students in the class, two groups can be distinguished: small lectures (LES) - a lecture class of 40 or fewer students, and large lectures (LEL) - a lecture class of more than 40 students. Since the number of MA students in Iranian universities does not exceed 40, we only had small lectures. These attributes along with the number of words and recording duration have been displayed in two separate tables for each corpus in Appendix A.

### *3.3. Data analysis*

In order to analyze data and find the frequency of the use of DMs in each lecture, WordSmith version 4.0, a concordance software program, was used. The variables used for the analysis of DMs were the number of occurrences and the frequency rate of each pre-established marker in the two sub-corpora. For analyzing the results of this study, the DMs classification model proposed by Bellés-Fortuño (2006) was selected. As mentioned before, his previous classification was unsuccessful because categories under micro and macro-markers did not comply with any firm linguistic rule. "... whereas some categories were miscellaneous, others were semantic, morpho-syntactic or even pragmatic" (Bellés-Fortuño, 2006, p. 156).

Therefore, Bellés-Fortuño (2006) developed a taxonomy of DMs based on the Hallidayan's (1994) functional meanings, namely, ideational, interpersonal and textual, and the relations they can convey along the discourse utterances. His model consists of three categories: micro- markers, macro-markers and operators.

"Micro-markers express logico-semantic relations in the discourse. According to this model, these markers have lexical or descriptive meaning" (Bellés-Fortuño, 2006; p. 95). Therefore, categories such as causal, contrastive, consecutive or additional DMs would be placed under this part.

"Macro-markers convey an overall structure of the ongoing discourse and aim at segmenting and structuring utterances. They enhance retention and recall in post-lecture tests" (Chaudron & Richards, 1986, p. 43). They play an essential role in activating content schemata (DeCarrico & Nattinger, 1988) and helping listeners to successfully follow the lecture (Khuwaileh, 1999).

"Operators are those DMs which rhetorically signal the speakers' intentions and affect the illocutionary force. These markers are more specifically related to conversational, spoken discourse rather than written discourse" (Lorente, 1996, p. 38). These have been traditionally called in the literature 'pragmatic markers'.

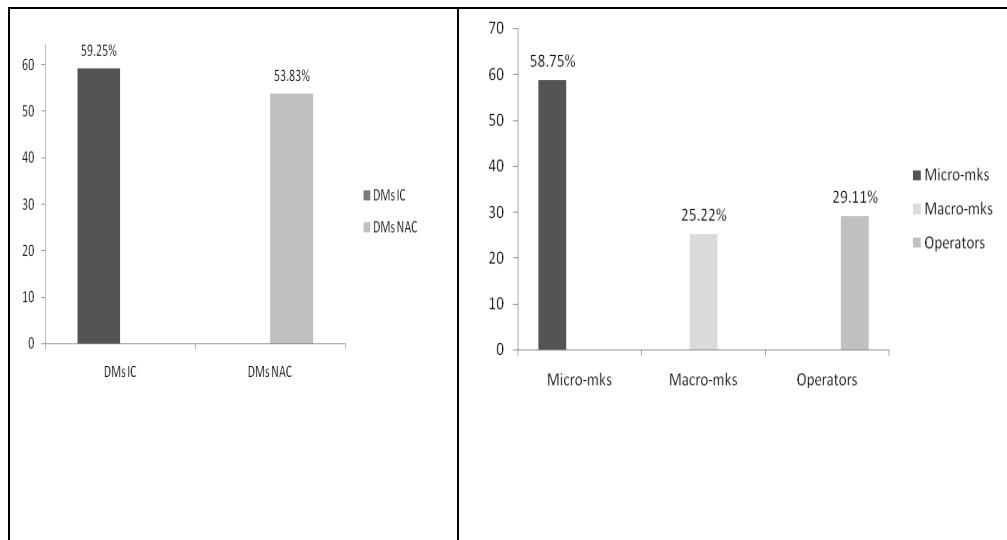


Because of the complexity and variety of DMs, Bellés-Fortuño (2006) narrowed the scope of his study by choosing a maximum of three DMs for each category. The resulting scheme consists of five different categories is shown in Appendix B.

## 4. Results

### 4.1. *The overall results of DMs*

The overall results from the analysis of DMs in the NAC and IC showed that DMs were used more in the IC than in the NAC (see Figure 1). The results obtained from both sub-corpora represented that micro-markers have got the highest rate followed by operators and macro-markers in the last instance (see Figure 2).



**Figure 1.**  
DM results in the NAC and IC

**Figure 2.**  
Total rate of micro-markers, macro-markers and operators in the NAC and IC

Regarding the frequency rate of micro-markers, the two sub-corpora did not differ significantly. The results of analyzing micro-markers in the NAC represented the three categories with the highest frequency rate, that is, Additional, Contrastive and Consecutive. In the IC, the most frequent categories were Additional and Contrastive followed by Temporal and Consecutive. As the results



displayed, Consecutive was the third category in the NAC whereas in the IC this category, representing a minor difference with Temporal, came in the fourth place. Categories showing the lowest frequency rate were Temporal and Causal in the NAC; however, in the IC, Causal represented the least frequency rate (see Table 1).

**Table 1.**

Micro-markers results in number of occurrences in the NAC and the IC

Micro-markers	in the NAC	in the IC
Additional	678	262
Temporal	119	83
Causal	106	55
Contrastive	276	110
Consecutive	189	83
Total	1.368	593

The three most recurrent micro-markers in the NAC were in order: *and*, *but* and *so* which were similar in the IC although with a different frequency rate. It worth mentioning that due to the extensive use of micro-marker *and*, the additional category showed the highest frequency rate. The major difference found in both sub-corpora was manifested by the Causal and the Contrastive categories in which the micro-markers *since* and *however* show no instance in the IC.

According to the findings, the total frequency rate of macro-markers in the IC was slightly higher than the total frequency rate in the NAC. The most frequent category in both sub-corpora was Topic-shifter due to the large number of occurrences of the macro-marker *so*. The next categories in order of frequency rate were Rephraser, Organizer and Starter in the NAC; however, Starter came in the third place in the IC followed by Organizer.

The least frequently used category was Conclusion in the NAC and the IC with this difference that no instance was found for the macro-markers of this category in the IC. The reason for the non-occurrence of these macro-markers was that all the lecturers in the IC code switched at the end of their classes and finished their lecture in their mother tongue. Another difference was that two other macro-markers, namely, *that is* and *to begin with*, represented no occurrence in the IC. Considering individual macro-markers, it has been found that in both sub-corpora the three most recurrent macro-markers were: *so*, *now* and *actually* in the Topic-shifter (see Table 2).



**Table 2.**

Macro-markers results in number of occurrences in the NAC and the IC

Macro-markers	in the NAC	in the IC
Starter	30	18
Rephraser	42	24
Organizer	32	11
Topic- shifter	360	235
Conclusion	17	0
Total	488	278

Operators revealed a higher frequency rate in the IC in contrast with the NAC. In the analysis of operators in the IC, the three most recurrent categories were found to be identical with those in the NAC which were Pause- filler, Confirmation-check and Acceptance, although with different frequency rate. In the IC, the most recurrent operators were *well* and *okay?*, whereas, in the NAC *and* followed by *okay* showed the highest rate.

Categories showing the lowest frequency rate were also similar in both sub-corpora, which were: Attitudinal and Elicitation. The least used category was Elicitation in both the NAC and the IC. Among these operators, the following ones revealed no instance in the IC, namely, *as you know*, *I/we believe* and *anyone?* (see Table 3).

**Table 3.**

Operator results in number of occurrences in the NAC and the IC

Operators	in the NAC	in the IC
Attitudinal	43	18
Pause filler	208	184
Elicitation	7	7
Acceptance	84	70
Confirmation	93	106
Total	435	385



#### 4.2. The overall collocations results

A closer look at the whole DMs collocations in the NAC and IC revealed that some specific DMs collocate with each others. The most recurrent collocations in both sub-corpora have been collected and analyzed. According to the results, the NAC used more collocations in contrast with the IC (see Tables 4 and 5). It was found that micro-markers tended to cluster with other markers more than macro-markers or operators. The reason could be due to the fixed and less variable structure of micro-markers, in contrast with the unsteady and less permanent structure of macro-markers.

**Table 4.**

DMs collocations in the NAC

DMs collocations in the NAC	# DMs	%
<i>and so</i>	57	1.3
<i>and then</i>	56	1.3
<i>okay? and</i>	13	0.3
<i>okay so</i>	9	0.2
<i>but then</i>	6	0.1
<i>and finally</i>	3	0.07
Total	84	3.27

**Table 5.**

DMs collocations in the IC

DMs collocations in the IC	# DMs	%
<i>and then</i>	30	1.5
<i>okay so</i>	11	0.5
<i>okay now</i>	7	0.3
<i>okay and</i>	4	0.2
<i>and now</i>	2	0.1
Total	54	2.6

According to the results obtained from the NAC, the DM *and* was found to collocate more than any other DMs. It frequently co-occurred with other markers such as *so*, *then*, *okay* and *finally*. The next marker which represented a high frequency rate regarding the NAC clusters was *okay* that collocate with markers such as *and* and *so*.

After analyzing the collocation examples it was found that in most cases when two DMs collocated, the first marker adopted a meaning closer to the next marker. For instance, in the collocation *and so*, the DM *and* conveyed a meaning closer to consecutive *so*. The similar instances were: *and then* and *okay so*.

However, in some cases both markers maintained their initial meaning. For example, when the operator *okay?* which functioned as confirmation-check, accompanied by the additional micro-marker *and*, they both kept their initial functions and meanings. The other case was *and finally* in which *and* maintained its additional meaning. The collocation *but then* was another instance in



which the marker *then* got both a temporal and a consecutive meaning based on the discourse context.

The results in the IC also illustrated that the markers *and* and *okay* showed the highest collocation among other markers. The identical collocation in both sub-corpora were *and then* and *okay so*. As opposed to the NAC, the IC represented other clusters, namely, *okay now*, *okay and* and *and now*. In the case of *okay now*, both markers functioned as a Topic-shifter; in other words, both markers maintained their initial meaning. The last cluster in the IC, *Okay and* and *and now*, also showed similar results.

## 5. Discussion

Previous studies mainly focused on the role of DMs in written discourse; however, little research has focused on the role and function of DMs in spoken academic discourse, especially in an Iranian contrastive EFL context. Therefore, this study aimed at comparing the two sub-corpora, the NAC and the IC, in order to analyze the use and functions of DMs and find the similarities and differences that might be encountered. In addition, this contrastive study looked for those DMs that normally co-occur in both the English and Iranian EFL lectures.

The results showed that, although the average length and word per lecture was higher in the NAC than in the IC, as opposed to the average time of lectures which was slightly higher in the NAC than the IC, DMs were more often used in the IC than in the NAC. Therefore, it could be mentioned that the use of DMs in general was equivalent to the average length and words per lecture in both sub-corpora. However, it did not mean that both sub-corpora had used DMs entirely identical.

An overall view on individual category showed that micro-markers have got the highest rate in contrast with macro-markers and operators. This represented that lecturers in both sub-corpora tended to express logico-semantic relations in their lectures. As a result, it could be said that in the discourse of lectures there was a need to convey lexical and descriptive meaning by using micro-markers. Operators came in the second place based on the frequency rate in both the NAC and IC. These markers express the speakers' intentions and affect the illocutionary force.

The last type of markers in this study was macro-markers in both the NAC and IC. This might be due to the peculiarities of spoken academic discourse. Bellés-Fortuño (2006) provided a good explanation for the scarce need for macro-markers. He said: "lecturers and students have a well-



developed knowledge of the structure and framing of a lecture and therefore macro-markers are less needed" (Bellés-Fortuño, 2006, p. 286).

The results of this study are found to be beneficial for both English and Iranian EFL lecturers who are interested in the academic education. In addition, English and Iranian lecturers can improve their lecture through developing conscious knowledge of the use and functions of DMs. Generally, the findings address university lecturers who want to improve their lecture discourse by using linguistic signaling devices in the correct way according to use, function, and cross-cultural differences. Students who want to improve their lecture comprehension process for their successful academic training and results can also benefit from the results of this contrastive study. Teachers and course designers can also use the findings of this corpus study. These findings indicate what linguistic and discourse features should be taught in the classrooms, and more certainly, what should be incorporated into the EFL courses.

### Acknowledgement

[Supported by National Elite Foundation]

Special thanks go to National Elite Foundation for all their sincere support.

### References

- Badger, R., White, G., Sutherland, P., & Haggisc, T. (2001). Note perfect an investigation of how students view taking notes in lectures. *System*, 29, 405-417. Retrieved March 16, 2009, from <http://www.sagepub.com>
- Bathia, V. K. (2002). A generic view of academic discourse. In J. Flowerdew (Ed.), *Academic discourse* (pp. 1-39). London: Longman Pearson Education.
- Bellés-Fortuño, B. (2006). Discourse markers within the university lecture genre: A contrastive study between Spanish and North-American lectures. Unpublished Doctoral dissertation. Universitat Jaume I. Retrieved October 6, 2009, from <http://www.sciencedirect.com>
- Bellés-Fortuño, B., & Fortanet, I. (2004). Handouts in conference presentations. In I. Fortanet, J.C. Palmer, and S. Posteguillo (Eds.), *Linguistic studies in academic and professional English* (pp. 63-76). Castelló: Publicacions de la Universitat Jaume I.
- Chaudron, C., & Richards, J. (1986). The effect of discourse markers on the comprehension of lectures. *Applied Linguistics*, 7(2), 113-127.
- Crawford, B. (2004). Interactive discourse structuring in L2 guest lectures: Some insights from a comparative corpus-based study. *Journal of English for Academic Purposes*, 3, 39-54.



- Del Saz, M. (2003). *An analysis of English discourse markers of reformulation*. Unpublished PhD dissertation. València: Universitat de València. Retrieved November 3, 2009, from <http://www.sagepub.com>
- Flowerdew, J. (2003). Register-specificity of signaling nouns in discourse. In P. Leistyna, & F. Meyer (Eds.), *Corpus analysis: Language and structure and language use* (pp. 35-46). Amsterdam: Editions Rodopi B.V.
- Fortanet, I. (2004). The use of 'we' in university lectures: References and function. *English for Specific Purposes*, 23(1), 45-66.
- Halliday, M. A. K. (1989). *Spoken and written language*. Oxford: Oxford University Press.
- Jung, E. H. (2003). The role of discourse signaling cues in second language listening comprehension. *The Modern Language Journal*, 87(4), 562-577.
- Lee, J. J. (2009). Size matters: an exploratory comparison of small- and large-class university lecture introductions. *English for Specific Purposes*, 28, 42-57. Retrieved June 4, 2009, from [www.sciencedirect.com](http://www.sciencedirect.com).
- Rendle-Short, J. (2003). So what does this show us? Analysis of the discourse marker so in seminar talk. *Australian Review of Applied Linguistics*, 23, 46-62.
- Schiffrin, D. (1987). *Discourse markers*. Cambridge: Cambridge University Press.
- Young, L. (1994). University lectures-macro-structure and micro-features. In Flowerdew, J. (Ed.), *Academic listening: Research perspectives* (pp. 159-176). Cambridge: Cambridge University Press.

## Appendix A

### a) American English lectures

	Title	Primary discourse mode	Speech event	Number of words	Recording duration
LE1	History of the American Family Lecture	Monologic lecture	Large lecture	10,621	81min
LE2	Media Impact in Communication Lecture	Monologic lecture	Large lecture	9,164	72min
LE3	Twentieth Century Arts Lecture	Monologic lecture	Large lecture	5,981	41min
LE4	Intro Psychology Lecture	Monologic lecture	Large lecture	7,266	47min



LE5	Intro to Psychopathology Lecture	Monologic lecture	small lecture	7,938	52min
Total				40,970w	293min
Average words & time				8,19w	58.6min

**b) Iranian EFL lectures**

	Title	Primary discourse mode	Speech event	Number of words	Recording duration
LE1	Operational and theoretical definitions of variables	Monologic lecture	small lecture	3,530	91 min
LE2	Chaos complexity theory	Monologic lecture	Small lecture	3,457	41 min
LE3	Social constrains of speech	Monologic lecture	Small lecture	3,623	65 min
LE4	Rationalists and empiricists	Monologic lecture	Small lecture	5,421	71 min
LE5	Data analysis of discourse	Monologic lecture	Small lecture	3,941	55 min
Total				19,972w	323 min
Average words & time				3,99w	64.6 min

## Appendix B

**a) Micro markers**

Additional	Temporal	Causal	Contrastive	Consecutive
<i>and</i>	<i>Then</i>	<i>because</i>	<i>but</i>	<i>So</i>
<i>or</i>	<i>After</i>	<i>since</i>	<i>although/ though/ even though</i>	<i>Then</i>
<i>now</i>	<i>Before</i>	<i>because of</i>	<i>however</i>	<i>so that</i>



b) Macro markers

Starter	Rephraser	Organizer	Topic- shifter	Conclusion
<i>first (of all)</i>	<i>I mean</i>	<i>let's (let us) try, go back/through focus, look</i>	<i>so</i>	<i>Finally</i>
<i>to begin (with), we're gonna begin, let's begin</i>	<i>in other words</i>	<i>let me (lemme) go back/through focus, look</i>	<i>now</i>	<i>to end up/ with, to finish/up</i>
<i>I want to/ wanna do today/start with/talk about</i>	<i>that is</i>	<i>I wanna/want to discuss, do, emphasize...</i>	<i>actually</i>	<i>I'll see you</i>

c) Operators

Attitudinal	Pause filler	Elicitation	Acceptance	Confirmation
<i>I think/we think</i>	<i>And</i>	<i>any questions(?)</i>	<i>okay</i>	<i>okay?</i>
<i>as you know</i>	<i>Well</i>	<i>why is that?</i>	<i>alright</i>	<i>right?</i>
<i>I believe/we believe</i>	<i>Okay</i>	<i>anyone?</i>	<i>right</i>	<i>alright?</i>

## Bio Data

**Zohre Eslami-Rasekh** is associate professor of ESL Education in the Department of Teaching, Learning, and Culture in the College of Education at Texas A&M University in College Station. She has taught and educated teachers of ESL/EFL for more than 10 years. Her current research interest include sociocultural aspects of ESL teaching and assessment, intercultural and developmental pragmatics, and ESL teacher education.

**Abbas Eslami-Rasekh** is associate professor in the Faculty of Foreign Languages at Isfahan University in Iran. He has taught graduate and undergraduate level courses in English as a second language for the last 10 years. His research interests include issues related to the education of EFL and EAP students, EAP needs assessment.



**Shahla Simin** is a PhD candidate in Applied Linguistics and TEFL. She received her BA and MA in TEFL from University of Isfahan, Iran. She has over 15 years of teaching experiences at universities and language centers. She had published several research papers in journals such as "Comparative Literature and Culture" (ISI journal), "ESP World Journal", Asian EFL Journal, International Journal of Language Studies. Her research interests lie in pragmatics, sociopragmatics, discourse analysis, semiotics, sociolinguistics and language teaching.



## **Opening up a New Pathway to the Future of ELT: Construction and Validation of a Scale for the Measurement of L2 Teachers' Life-Responsive Language Teaching Perceptions**

**Reza Pishghadam<sup>1</sup>**

*Ferdowsi University of Mashhad*

**Reza Zabihi**

*University of Isfahan*

**Momene Ghadiri**

*University of Isfahan*

### **Abstract**

It is widely acknowledged that learning life skills is a fruitful practice that helps individuals to deal effectively with everyday challenges of life. With the advent of Applied ELT as a new paradigm in second/foreign language studies, ELT has gained a new and contributory status among other disciplines (Pishghadam, 2011). In this connection, it seems that due to their unique features which can provide ample opportunities for enhancing learners' life qualities as well as satisfying their educational needs, ELT classes can be sites where these life skills should be dealt with and improved more than any other educational context, hence the need for incorporating life skills training into the ELT curriculum in the form of Life Syllabus (Pishghadam & Zabihi, 2012). As one of the first attempts at eliciting language teachers' beliefs about the incorporation of life skills training into the ELT curriculum, the present study was conducted to construct a Life-Responsive Language Teaching Beliefs questionnaire in the foreign language context of Iran, making use of the defining indexes of life skills education proposed by the World Health Organization (WHO). Given the meritorious factorability of the present sample, the scale was then validated by conducting the Exploratory Factor Analysis (EFA). The results of several EFA analyses revealed that a four-factor solution, i.e. (a) Life-wise Empowerment, (b) Adaptability Enhancement, (c) Pro-social Development and (d) Life-over-language Preference, would yield a parsimonious grouping of the items. Finally, some pathways to future research were mapped out for other interested researchers to give Applied ELT more empirical adequacy.

**Keywords:** Applied ELT, Life skills training, Life-responsive language teaching perceptions, WHO, Life syllabus

---

<sup>1</sup> Email: rpishghadam@yahoo.com



## 1. Introduction

The underlying assumption of humanistic education is that education should empower individuals to lead a good life. In this connection, there would be “only one subject-matter for education and that is life in all its manifestations” (Whitehead, 1929, p. 6). This is not to suggest that attention to content area should be waived for the sake of improving students’ lives; on the contrary, it states that teaching should not be confined to facilitate students’ achievement of academic excellence but, as several educational philosophers such as Dewey (1897), Freire (1998), Krishnamurti (1981), and Walters (1997) have turned their focal attention to the incorporation of life issues in education, it should also consider issues of prime importance in students’ lives such as their emotions, relationships, thinking styles, attitudes, values and feelings. Be that as it may, several scholars have considered life issues such as, *inter alia*, happiness (Noddings, 2003), emotional abilities (Matthews, 2006), self-determination (Walker, 1999), autonomy (Winch, 1999) and critical thinking (Hare, 1999) as educational objectives to be attained.

Learning life skills is thus a fruitful practice (Murthy & Wig, 2003) that helps individuals to deal effectively with everyday challenges of life (Orley, 1997); accordingly, life skills training can enable students to act in pro-social ways (Birell Weisen & Orley, 1996) and help them take more responsibility for their behaviors and actions (Orley, 1997). Given the general consensus as to the importance of including life skills training in the context of education (Francis, 2007; Goody, 2001; Larson & Cook, 1985; Radja, Hoffmann, & Bakhshi, 2008; Spence, 2003), and by virtue of their high credibility with parents and community members (WHO, 1997), schools can be regarded as indispensable places where several life skills can be worked on alongside other academic subjects (Behura, 2012; Matheson & Grosvenor, 1999).

One of the forerunners of life skills education is World Health Organization (WHO) which has primarily been established with the aim of enhancing children’s mental and social well-being. In this view, life skills are defined as “the abilities for adaptive and positive behavior that enable individuals to deal effectively with the demands and challenges of everyday life” (WHO). The pivotal life skills emphasized by WHO include psychosocial and interpersonal competencies such as, decision making, problem solving, creative thinking, critical thinking, effective communication, interpersonal relationship skills, self-awareness, empathy and understanding, coping with emotions, and coping with stress.

In a similar vein, the United Nations International Children’s Emergency Fund (UNICEF) has also defined life skills as “a behavior change or behavior development approach designed to address a balance of three areas: knowledge, attitude and skills” ([www.unicef.org](http://www.unicef.org)). In this respect, the Delors



Report (Delors et al., 1996), whose mission is to give education the role of providing humanity with the capacity to control its own development, was put forward with four educational pillars, namely *learning to be, learning to know, learning to live together, and learning to do*. These advances have led to the preparation of a proposal entitled Education for Human Development which is based on the idea that any education has the responsibility to generate learning as well as to help students develop their other potentials and capabilities. Attempts have accordingly been made by some organizations such as UNESCO and the Ayrton Senna Institute to apply the four fundamental areas of learning proposed by Delors et al. (1996) with the aim of catering for and nourishing different aspects of individuals' lives such as their multiple competencies, abilities, innate potentials, as well as their emotions and attitudes.

It seems that the same argument can be put forward for the field of English Language Teaching (ELT). In order to get to grips with this, one needs to clarify what is really meant by the term *Applied ELT* (Pishghadam, 2011). Applied linguistics was once considered to be synonymous with language teaching (Strevens, 1992). In the course of time, ELT was being studied as a part of applied linguistics, and is currently highly open to receive ideas from several branches of applied linguistics such as sociolinguistics and psycholinguistics. These areas have enlightened the English language teaching professionals in enhancing their understanding of the field. Yet it is widely acknowledged that these ideas are in the form of recipes for enhancing the field of English language teaching and learning. It has been argued that these prescriptions, however, turn ELT practitioners into consumers of the findings of other domains of knowledge (Schmitt, 2002).

Given that any discipline has two complementary sides of theoretical and applied, it seems that, in the case of the field of ELT, the applied part has been ignored. In this connection, Pishghadam has argued that language teachers should no longer be consumers of the findings of other disciplines; rather, they should try to take on a more contributory and life-changing role. That is to say, ELT, by virtue of its interdisciplinary and scientific propriety, has now obtained the capability to be applied and contribute to other domains of knowledge. In addition, by virtue of their distinctive features, ELT classes can be the sites where several life skills may be prescheduled to be enhanced (Pishghadam, 2011). Among the unique features of ELT classes are (a) discussing a large number of social, scientific, and political topics, (b) holding pair work and group work in class, (c) comparing two cultures, (d) getting acquainted with the words and grammar of another language, (e) speaking in another language in which one can show their own real self, (f) taking language learning very seriously, and (g) having a funny and friendly atmosphere for learning. The idea was furthered by Pishghadam and Zabihi (2012) whose aim was to introduce *Life Syllabus* in order to encourage the ELT community to consider the enhancement of several life skills prior to language learning.



In the light of this corollary, ELT classes can thus be magnificent places for a life skills intervention. In this manner, these classes can act as a safety net, by teaching learners a variety of life skills. These distinctive features sets the scene for the corroboration of the Applied ELT theory wherein English teaching practitioners are expected to enhance different life skills such as critical thinking, empathy, creativity, motivation, stress management, etc. alongside teaching a language. Therefore, in this study we attempted to examine whether and to what extent Iranian EFL teachers believe that life skills should be incorporated into the ELT curriculum. To this end, a questionnaire was designed and validated to investigate language teachers' perceptions about life-responsive language teaching in the Iranian EFL context. The following research questions were addressed in this study:

1. What are the underlying factors of the life-responsive language teaching beliefs scale?
2. What is the nature of EFL teachers' life-responsive language teaching beliefs?

## 2. Method

### 2.1. Participants

A total of 192 Iranian EFL teachers at the University of Isfahan, Sheikhbahaei University, Islamic Azad University (Khorasgan Branch) and Abadan Institute of Technology (AIT) participated in this study. They ranged in age from 24 to 55 and were graduate-level university professors of English teaching, translation, linguistics and literature. The sample consisted of both male and female EFL teachers. The number of male and female teachers was 73 and 119 respectively. The detailed information on the participants is shown in Table 1.

**Table 1.**  
Distribution of participants

Sex	Age	Degree		Teaching Experience		
Male	73	Twenties	149	English Teaching	77	Less than two years
Female	119	Thirties	28	English Literature	46	2-4 years
		Forties	8	English Translation	62	5-9 years
		Fifties	7	English Linguistics	7	10-14 years
						15-19 years
						20-24 years
						25-30 years
Total	192					192



## *2.2. Data collection and analysis*

The first phase of the study involved the development and validation of a Life-Responsive Language Teaching Beliefs questionnaire, making use of the defining indices of life skills education proposed by the World Health Organization (WHO) and validating it by conducting the Exploratory Factor Analysis (EFA). This resulted in a checklist of life-responsive teaching indicators based on which the items of the questionnaire were constructed. Joint consultations were held to make the best items out of the previously made checklist and also to revise the items. Moreover, having obtained the opinions of six specialists in the Sociology of Education (2), Psychology of Education (3), and Applied Linguistics (1), the researchers revised the newly-designed questionnaire, accordingly.

Besides, fifteen EFL teachers whose characteristics were similar to those of the target population were asked to read the revised scale and to express their understanding of it. In doing so, the researchers were able to remove any ambiguity which could have otherwise led to misunderstandings on the part of the target participants. The questionnaire was then administered to 192 EFL writing teachers in two ways: (a) face-to-face administration (print form), and (b) administration via an electronic fillable PDF document. The questionnaire comprised 45 items that were scored according to the Likert-type scale of seven points ranging from (1) "strongly disagree" to (7) "strongly agree". Besides, the participants were asked to provide answers to questions regarding their age, gender, major, degree, and teaching experience.

Next, the internal consistency of the whole questionnaire was examined with the Cronbach Alpha reliability estimate. Moreover, using Cronbach Alpha, the reliability of each factor constructing the validated questionnaire was also measured. Further, to validate the questionnaire, EFA was used. First, The Factorability of the intercorrelation matrix was measured by two tests: Kaiser-Meyer-Olkin test of Sampling Adequacy (KMO) and Bartlett's Test of Sphericity. The results obtained from the two tests revealed that the factor model was appropriate. Then, PCA extracted the underlying factors by calculating the eigenvalues of the matrix greater than 1.0. The Scree test was utilized for the purpose of deciding about the number of factors to retain for rotation. For conducting factor rotation, Varimax (orthogonal rotation) with Kaiser Criterion was used. The result was a rotated component matrix and a transformation matrix. The rotated component matrix indicated the variables loaded on each factor so that the researchers came up with the new factors.



### 3. Results

#### 3.1. Reliability of the life-responsive language teaching beliefs questionnaire

Cronbach Alpha estimated the reliability of the whole scale as 0.94. All of the four factors yielded good reliability estimates ranging from 0.57 to 0.90 (Table 2).

**Table 2.**

Reliability of each factor

Factors	Cronbach's Alpha	N of Items
Factor 1	.907	15
Factor 2	.881	14
Factor 3	.840	10
Factor 4	.571	6
Total	.941	45

#### 3.2. Construct validation of the newly designed scale

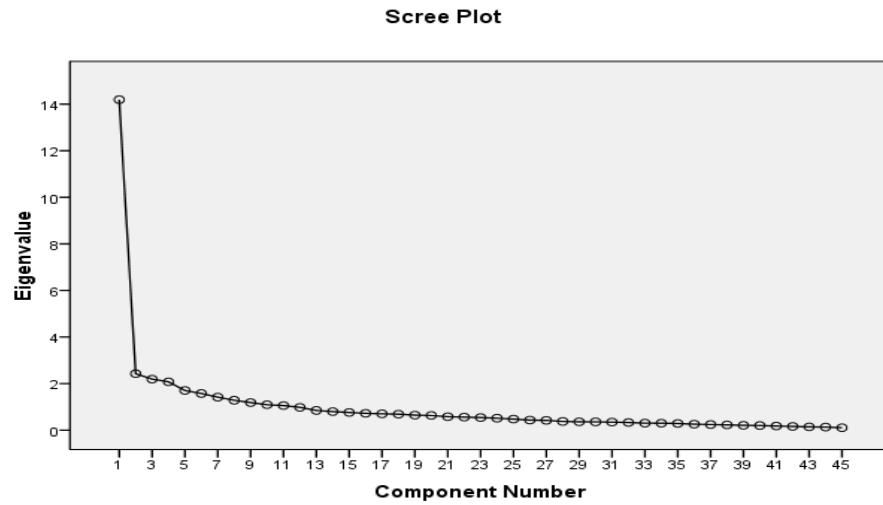
The Factorability of the intercorrelation matrix was assessed by two tests: Kaiser-Meyer-Olkin test of Sampling Adequacy (KMO) and Bartlett's Test of Sphericity. The KMO statistic obtained in this study was .88. As Kaiser (1974) has pointed out, KMOs in the .80s is "meritorious", hence the appropriateness of the selected sample for factorial analysis (Table 3).

**Table 3.**

KMO and Bartlett's test

KMO and Bartlett's Test		
Kaiser-Meyer-Olkin Measure of Sampling Adequacy		.889
Bartlett's Test of Sphericity	Approx. Chi-Square	3.814E3
	Df	990
	Sig.	.000

The construct validity of *Life-Responsive Teaching Beliefs Questionnaire* was examined through EFA. PCA extracted 11 factors with eigenvalues greater than 1.0 which accounted for 67% of the variance. The results obtained from the Scree Test, however, indicated that a four-factor solution might provide a more suitable grouping of the items in the questionnaire (Figure 1).



**Figure 1.**

The Scree test

The researchers, then, inspected orthogonal rotation. The result of Varimax with Kaiser Normalization was a rotated component matrix. The results of this analysis are shown in Table 4. The results indicated that the first factor consisted of 15 items; the second factor consisted of 14 items; factor 3 consisted of 10 items; and factor 4 consisted of 6 items. The total number of items was 45.

**Table 4.**

Rotated components obtained via principal component analysis and their loadings

Component 1	Component 2	Component 3	Component 4
25= .71	9= .80	14= .65	7= .74
26= .71	10= .73	41= .61	18= .60
31= .67	2= .59	6= .58	32= .51
33= .62	11= .55	24= .54	21= .41
23= .60	39= .54	4= .52	44=.39
27= .54	40= .53	15= .51	1=.25
29= .54	5= .50	37= .51	
45= .53	12= .50	35= .49	
38= .53	13= .49	42= .39	
30= .52	3= .48	8= .36	



28= .52	16=.39		
43= .52	20=.31		
36= .51	17=.31		
34= .50			
22= .33			

Rotation Method: Varimax with Kaiser Normalization

Finally, the researchers analyzed the items comprising each factor and named the four factors as (i) *Life-wise Empowerment*, (ii) *Adaptability Enhancement*, (iii) *Pro-social Development*, and (iv) *Life-over-language Preference*. Items representing each factor are displayed in Table 5, and the validated questionnaire is given in Appendix 1.

**Table 5.**

Five factors of the questionnaire

# Areas	Statements	N of items	Percentage
1. <i>Life-wise Empowerment</i>	25, 26, 31, 33, 23, 27, 29, 45, 38, 30, 28, 43, 36, 34, 22	15	33.33
2. <i>Adaptability Enhancement</i>	9, 10, 19, 2, 11, 39, 40, 5, 12, 13, 3, 16, 20, 17	14	31.11
3. <i>Pro-social Development</i>	14, 41, 6, 24, 4, 15, 37, 35, 42, 8	10	22.22
4. <i>Life-over-language Preference</i>	7, 18, 32, 21, 44, 1	6	13.33
	Total	45	100

### 3.3. Descriptive statistics of the questionnaire data

The results from descriptive statistics revealed that EFL teachers in these four universities put more emphasis on the Pro-social Development of language learners (mean: 5.49). Since the questionnaire was 7-scale Likert, the maximum and minimum scores were 7×1 respectively. As it is displayed in Table 6, teachers lent prominence to different factors of life-responsive teaching from the most to the least as follows: 1) *Pro-social Development* (mean: 5.49), 2) *Life-wise Empowerment* (mean: 4.98), 3) *Adaptability Enhancement* (mean: 4.93) and *Life-over-language Preference* (mean: 4.21). It should be noted that the maximum score calculated was 7.



**Table 6.**

The means of factors obtained from descriptive statistics

	<b>N</b>	<b>Mean</b>	<b>Std. Deviation</b>
Pro-social Development	192	<b>5.49</b>	.76
Life-wise Empowerment	192	4.98	.83
Adaptability Enhancement	192	4.93	.83
Life-over-language Preference	192	4.21	.84
Valid N (listwise)	192		

#### 4. Discussion

Despite the high degree of prominence given to life skills training in the context of education (Francis, 2007; Goody, 2001; Larson & Cook, 1985; Radja, Hoffmann, & Bakhshi, 2008; Spence, 2003), it seems that there is a paucity of research examining the consideration of life issues in the field of English language teaching. This study sought to find out the reliability and validity of the Life-responsive Language Teaching Beliefs questionnaire when examined with Iranian EFL teachers. We have utilized EFA as an efficient tool for determining the underlying factors of the instrument. The results have revealed that a combination of four factors could best represent the underlying structure of life-responsive language teaching instrument.

Having analyzed the items of the questionnaire, the researchers tried to name each factor. These factors were labeled as: Life-wise empowerment, Adaptability enhancement, Pro-social development, and Life-over-language preference. The reasons for the selection of such terms are explained below. *Life-wise empowerment* is the label for the first factor which consists of 15 items. It refers to the language teacher's ability to support mental well-being and behavioral preparedness of learners including creative and critical thinking. The second factor is called *Adaptability Enhancement* which refers to the language teacher's ability to foster adaptive and positive behaviors that enable individuals to deal effectively with the demands and challenges of everyday life including problem-solving and decision making. This factor consists of 14 items. Factor 3 which is known as *Pro-social Development* relates to the language teacher's ability to promote personal and social development including interpersonal bonds and effective communication. This factor comprises 10 items. *Life-over-language Preference* is the label selected by the researchers for the fourth factor of the questionnaire which includes 6 items characterizing the language teacher's ability to center attention on learners' qualities of life including their feelings and emotions in comparison with linguistic points.



In addition, it was revealed that, among the four underlying factors of the questionnaire, the participants lent more prominence to the pro-social development of language learners. That is to say, they believed that language teachers should try to promote, prior to other skills, learners' personal and social development including interpersonal bonds and effective communication. Moreover, other underlying factors of life-responsive language teaching, i.e. life-wise empowerment, adaptability enhancement, and life-over-language preference, also received considerable attention by the participants.

We have made an important, novel contribution by commencing with attempts to operationalize the tenets of applied ELT (Pishghadam, 2011) and life syllabus (Pishghadam & Zabihi, 2012). That being said, there were some notable limitations to the present study: (i) First, because we did not conduct face-to-face interviews, we cannot confidently say to what extent L2 teachers' underestimating or underscoring the importance of life skills training in ELT classes has been due to their own will, their resistance or other institutional constraints; (ii) Second, our design did not allow us to assess whether L2 teachers' high- or low-life-responsiveness would have any role to play in their learners' improved or impaired satisfaction with life; (iii) Third, the current scale was designed based on the essential life skills proposed by only one well-known organization. We may have observed a different factor structure if we had used other models such as the Targeting Life Skills (TLS) Model which encompasses 35 life skills that have recurrently emerged as being essential for individuals to reach their full potential and lead a successful life (Hendricks, 1998).

Taken as a whole, the value of this research lies in designing a scale for the measurement of L2 teachers' life-responsive language teaching beliefs, making use of the pivotal life skills emphasized by the World Health Organization which include ten psychosocial and interpersonal competencies. Besides, the present scale can help researchers and practitioners interested in applied ELT to employ the scale in the field of language teaching and shed more light on the extent to which L2 teachers believe that the key life skills proposed by WHO should be incorporated into the ELT curriculum.

We hope that future research will lead to further evaluation and enhancement of this scale. Researchers should continue to carry out thorough assessment of the psychometric properties of the instrument designed to measure life-responsive language teaching beliefs. Only after the true factor structure of the life-responsive language teaching beliefs scale has been determined can researchers confidently assert conclusions about the perceived prominence that should be given to key life skills in ELT classes. Moreover, replications of this study on different samples and with different scales are essential in coming up with a solid understanding of English teachers' perceptions of life-responsive teaching. Besides, ELT researchers and practitioners should try to



speculate about the meaning of life skills training in their own context, and to consider how they should promote life skills in their learners. The key to achieving this goal is for teacher educators to help language teachers become *Educational Language Teachers*, i.e. “critical, forethoughtful, and interdisciplinary educators who in addition to having a thorough command of language have a fair knowledge of other disciplines” (Pishghadam, Zabihi, & Kermanshahi, 2012, p. 896).

Further research will also be required to determine the effect of life-responsive language teaching perceptions on learners' satisfaction with life and whether the higher reporting of learners' satisfaction with life are the result of higher mean scores of the four factors of the current scale. In the end, researchers are recommended to examine objectively the relationship between life-responsive language teaching perceptions and other related variables such as teacher personality, autonomy, self-efficacy, and burnout. Finding any association between these variables and life-responsive language teaching perceptions helps us have a better understanding of the role of these psychological factors in making L2 teachers respond to applied ELT.

## References

- Behura, S. (2012). A review on life skills education in schools. *Elixir Psychology*, 43, 6790-6794.
- Delors, J. et al. (1996). Learning: The treasure within, report to UNESCO of the international commission on education for the twenty-first century. UNESCO, Paris.
- Dewey, J. (1897). My pedagogic creed. *The School Journal*, 54(3), 77-80.
- Francis, M. (2007). Life skills education. [www.changingminds.org](http://www.changingminds.org).
- Freire, P. (1998). *Pedagogy of freedom: Ethics, democracy, and civic courage*. Lanham, Maryland: Rowman & Littlefield Publishers, Inc.
- Goody, J. (2001). Competencies and education: Contextual diversity. In: D. S. Rychen, & L.H. Salganik (Eds.), *Defining and selecting key competencies*. Gottingen, Hogrefe and Huber Publications.
- Hare, W. (1999). Critical thinking as an aim of education. In R. Marples (Ed.), *The aims of education*. London: Routledge.
- Hendricks, P. (1998). Developing youth curriculum using the targeting life skills model: Incorporating developmentally appropriate learning opportunities to assess impact of life skill development (Tech. Rep. No. 4H-137A). Ames: Iowa State University Extension.
- Kaiser, H. F. (1974). An index of factorial simplicity. *Psychometrika*, 39(1), 31-36.
- Krishnamurti, J. (1981). *Education and the significance of life*. HarperCollins Publishers.
- Larson, D. G., & Cook, R. E. (1985). Life-skills training in education. *Journal of Group Psychotherapy, Psychodrama, & Sociometry*, 38(1), 11-22.



- Matheson, D., & Grosvenor, I. (Eds.) (1999). *An Introduction to the study of education*. London: David Fulton Publisher.
- Matthews, B. (2006). *Engaging education: Developing emotional literacy, equity, and co-education*. McGraw-Hill Education: Open University Press.
- Murty, S., & Wig, W. (2003). Who bothers about mental health care? *The Tribune* (24.12.2003).
- Noddings, N. (2003). *Happiness and education*. Cambridge: Cambridge University Press.
- Orley, J. (1997) Promoting Mental Health and Teaching Skills for life: The WHO Approach. [online] [www.healthchildrennetwork.ln/pdf/conference/1997/orleyenpdf.2003](http://www.healthchildrennetwork.ln/pdf/conference/1997/orleyenpdf.2003).
- Pishghadam, R. (2011). Introducing applied ELT as a new approach in second/foreign language studies. *Iranian EFL Journal*, 7(2), 9-20.
- Pishghadam, R., & Zabihi, R. (2012). Life syllabus: A new research agenda in English language teaching. *TESOL Arabia Perspectives*, 19(1), 23-27.
- Pishghadam, R., Zabihi, R., & Kermanshahi, P. (2012). Educational language teaching: A new movement beyond reflective/critical teacher. *Life Science Journal*, 9(1), 892-899.
- Radja, K., Hoffmann, A. M., & Bakhshi, P. (2008). *Education and capabilities approach: Life skills education as a bridge to human capabilities*. [http://ethique.perso.neuf.fr/Hoffmann\\_Radja\\_Bakhshi.pdf](http://ethique.perso.neuf.fr/Hoffmann_Radja_Bakhshi.pdf).
- Schmitt, N. (2002). *An introduction to applied linguistics*. London: Arnold.
- Spence, S. H. (2003). Social skills training with children and young people: Theory, evidence and practice. *Child and Adolescent Mental Health*, 8(2), 84-96.
- Strevens, P. (1992). Applied linguistics: An overview. In W. Grabe & R. B. Kaplan (Eds.), *Introduction to applied linguistics*. (pp. 13-31). Addison-Weley Publishing Company.
- Walker, J. C. (1999). Self-determination as an educational aim. In R. Marples (Ed.), *The aims of education* (pp. 112-123). London: Routledge.
- Walters, J. D. (1997). *Education for life: Preparing children to meet the challenges*. Crystal Clarity Publishers.
- Birell Weisen, R., & Orley, J. (1996). Life skills education: Planning for research as an integral part of life skills education development, implementation and maintenance. Geneva: WHO, Programme on Mental Health.
- Whitehead, A. N. (1929). *The aims of education and other essays*. New York: Macmillan.
- Winch, C. (1999). Autonomy as an educational aim. In R. Marples (Ed.), *The aims of education* (pp. 74-84). London: Routledge.
- World Health Organization (WHO). (1997). Life skills education for children and adolescents in schools: Introduction and guidelines to facilitate the development and implementation of life skills programmes. Geneva, Switzerland: WHO Programme on Mental Health.



## Appendix

### The life-responsive language teaching beliefs questionnaire

No.	Statement
1	The primary goal of language classes should be language learning.
2	Language teachers are responsible for enhancing learners' well-being.
3	The learners' ability to keep, make and end friendly relationships constructively should be practiced and enhanced inside English classes.
4	Language classes can become sites where learners' lives are enhanced.
5	Language classes can teach learners how to empathize with others.
6	Language teachers should have knowledge of other disciplines.
7	Attending to learners' feelings and emotions deviates language teachers from teaching the language.
8	Language classes have unique potentials for enhancing the quality of learners' lives.
9	An optimal goal of language learning tasks should be to enhance learner's decision-making capabilities.
10	An optimal goal of language learning tasks should be to enhance learner's problem-solving abilities.
11	Language classes can teach children the essentials of mental health by giving training in life skills.
12	Language classes should be sites to raise learners' self-awareness.
13	Coping with stress and emotions should be practiced through language learning tasks.
14	Language classes should be places where effective communication can be enhanced.
15	Language classes should be places where interpersonal bonds can be strengthened.
16	Language teachers should help learners build positive and healthy relationships.
17	Language classes should be sites where learner's cultural identity is constructed.
18	Life skills training in language classes may hinder the language learning process.
19	Learning how to relax, so that tensions created by unavoidable stress do not give rise to health problems should be the purpose of a language class.
20	Audio-visual activities in my class mainly deal with enhancing effective communication or interpersonal relationships.
21	Most feedbacks and discussions in my class revolve around linguistic points rather than enhancing the quality of learners' lives.
22	Language tasks should be designed in a way to enable learners to think diversely and extensively.
23	Language education should set out to facilitate the practice and reinforcement of learners'



	psychosocial skills in a culturally and developmentally appropriate way.
<b>24</b>	Language education can help the promotion of personal and social development.
<b>25</b>	Language education can help the prevention of health and social problems.
<b>26</b>	Language education can help the protection of human rights.
<b>27</b>	Language education can help other disciplines such as sociology, anthropology, psychology, psychotherapy, etc. to enrich themselves.
<b>28</b>	Supporting mental well-being and behavioral preparedness should be the primary purpose of language teaching.
<b>29</b>	Language teaching should equip children to behave in pro-social ways.
<b>30</b>	Language teaching should be regarded as a means of empowering children and enabling them to take more responsibility for their actions.
<b>31</b>	Language teachers should educate learners about life skills so that they can face with each trauma of life with greater resilience and higher efficiency.
<b>32</b>	Life skills education should only be the purpose of mental and behavioral disorder courses.
<b>33</b>	Effective application of language syllabi in English classes can influence the way children feel about others and themselves.
<b>34</b>	Effective application of language syllabi in English classes can contribute to learners' self-confidence and self-esteem.
<b>35</b>	English language institutes should gain a high credibility with parents and community members.
<b>36</b>	A comprehensive language teacher-training program is needed to produce teachers who can enhance learners' life qualities.
<b>37</b>	A comprehensive language teacher-training program is needed to produce teachers who can support children's educational and mental health requirements.
<b>38</b>	Language classes can act as a safety net by teaching students various life skills.
<b>39</b>	Fostering abilities for adoptive and positive behaviors that enable individuals to deal effectively with the demands and challenges of everyday life should be practiced in English classes.
<b>40</b>	Developing adaptive and forward-looking people should be the purpose of English education.
<b>41</b>	Active decision-making by assessing different options and what effects different decisions may have should be the purpose of language learning tasks.
<b>42</b>	Language learning tasks should be designed in a way that increase the learners' ability to perceive problems, seek and assess choices, make sound decision of each choice and solution and implement suitable solution.
<b>43</b>	Divergent thinking, i.e. thinking in different dimensions for new inventions, should be



	practiced through the proper design and application of language learning tasks.
<b>44</b>	The ability to recognize factors that influence attitudes and behavior such as values, peer pressure, and the media should be practiced through language learning tasks.
<b>45</b>	Language learners should become enabled to express their opinions and desires, needs and fears.

### Bio Data

**Reza Pishghadam** has a Ph.D. in Teaching English as a Foreign Language (TEFL) from Allameh Tabatabai' University in Tehran. He is currently on the English faculty of Ferdowsi University of Mashhad, Iran. He is now associate professor of TEFL, who teaches first language theories, socio-psychological aspects of language education, and applied linguistics. Over the last five years, he has published more than one hundred articles and books in different domains of English language education. In 2007, he was selected to become a member of the National Association of Elites of Iran. In 2010, he was classified as the top researcher in humanities by the Ministry of Sciences, Research, and Technology of Iran. His current research interests are: Psychology and Sociology of language education, Cultural studies, Syllabus design, and Language testing.

**Reza Zabihi** is a PhD candidate of TEFL at the University of Isfahan, Isfahan, Iran. He is also a member of the National Association of Elites of Iran. His major research interests include syllabus design as well as sociolinguistic and psycholinguistic studies. He has published 25 research articles in international journals and is currently teaching at University of Isfahan, Iran.

**Momene Ghadiri** is a PhD candidate of TEFL at the University of Isfahan, Isfahan, Iran. Her main areas of research are English teaching, discourse analysis, and sociolinguistics. She holds an MA degree in TEFL.



## Educational Therapy: On the Importance of Second Language Communication on Overcoming Depression

Azar Hosseini Fatemi

Safoora Navari<sup>1</sup>

*Ferdowsi University of Mashhad*

### Abstract

The notion that language learners bring with themselves cognitive and affective barriers to learning situation has long been acknowledged by different researchers. Many, especially in the field of Cognitive Psychology, have tried to find ways in order to alleviate such problems. However, language learners may carry the emotional problems that can be treated by teachers and by engaging in enjoyable activities in a second language learning environment (Naditz, 2007). Such glance at second language classroom is well understood in the notions of Psychodynamic psychology that emphasizes the importance of emotional experiences. Language learning thus can be a suitable approach for employing the psychodynamic therapeutic techniques to relieve various emotional problems of language learners under the realm of Educational therapy (Jarvis, 2005). Developed by Caspary in 1970s, Educational therapy makes use of psychodynamic principles and attempts to combine the psychoanalytic theory with teaching. Although this idea is supposed to be applied in the sessions of therapy, language teachers as some qualified therapists can devote more time and attention to emotionally problematic students and by highlighting the benefits of learning and communicating in a second language help those learners to overcome their emotional blocks such as depression. Many depressed learners may suffer from the senses of alienation and self-separation, however; it is thought that after spending some time in a second language learning classroom in which the teacher can probe into the emotional problems of a learner, depression state can be altered to that one of motivation to communicate in another language. This opening of unexpected doors to happiness that can occur through communicating in another language may not only help overcome the symptoms of depression but also lead to personal growth and a sense of fulfillment. Thus, it is most likely that following the psychodynamic principles of educational therapy, language teachers can turn the attention of depression suffering learners to the more beneficial and flourishing aspects of communicating in a second language.

**Keywords:** Communication, Depression, Educational therapy, Psychodynamic psychology

---

<sup>1</sup> Email: safoora.navari@gmail.com



## 1. Introduction

Education in itself is more than just dealing with teachers transmitting information to a group of learners and the issues related to that. It can also be considered as a more critical process containing the dynamic interplay between the teachers and learners as well as dealing with their individualistic personalities alongside the vital exchange of information. This crucial role of education can be well investigated in the realm of Educational psychology in which the focus lies in the application of psychology in education so much so that the principles of learning one special field can contribute to the notion of lifelong learning (Kaplan, 1990). Learning another language as a challenging and interesting process can also be recognized as one of those subjects whose learning can lead to a life-wise activity which can cause the inner satisfaction as well as the self-promotion of the individuals, even the troubled ones.

However, the previous approaches to education such as that of educational psychology mostly valued the significance of other majors like psychology in creating and monitoring the foreign language teaching theories. Pishghadam (2011) has recently introduced are a rather distinctive and rather challenging concept in ELT paradigm. In fact, with a new look to the issue of education, the notions of 'Applied ELT' and 'Life Syllabus' have opened a new window to the world of ELT, promising to deal with the more problematic issues in language teaching which have not been tapped before properly.

In this theory, it is believed that some life quality factors can be tackled mainly in this new look towards education. To put it in a nutshell, education here is seen as one area in which life qualities such as social relations, safety, physical health, emotional abilities and happiness can be triggered to improve (Pishghadam & Zabihi, 2012). Therefore, it can be assertive that both educational psychology and language teaching move under the flag of Humanism that holds, "there is only one subject-matter for education, and that is life in all its manifestations" (Whitehead, 1929, p. 6). As a matter of fact, it was upon this major school of thought- Humanism- that the special attention to the inner world of the learners and the emphasis on the feelings and emotions of the learners came on board (Williams & Burden, 2002).

Learners with special difficulties usually face more problems in the classroom that differ from those of others. The type of education that can help such troubled learners to benefit the learning atmosphere of the classroom as well as to become armed with capabilities to overcome their more serious problems in life even after the learning is over, is much of a compelling need these days. Thus, considering the importance of such problematic issues of learners in the context of education,



Life Syllabus's promise to concern with the more life-survival issues that can be induced through language learning, seems a considerable one which is worth analyzing.

## **2. Applied ELT and Life Syllabus**

In the continuum of the notions dealing with education, one area has been recently proposed as a new approach to ELT which alongside the introduction of 'Life Syllabus' is blooming in the field. In his theory, it was pointed out that the crucial role of specialists in language teaching field should shift from simply users to producers. By carefully pursuing the guidelines of applied theory, language teachers can make use of a 'life syllabus' which not only 'considers life issues as its first priority' (p. 8), but also in that syllabus' language learning is more purposeful and educational, revolving around more important issues of life' (Pishghadam, 2011).

By the same token, following the innovative guidelines of a 'life-syllabus' approach one can certainly hope to find teaching materials that can promise to foster a good command of English language on the one hand, and the inspiration of creative thought and action, which are among the vital prerequisites for a successful person in modern life, on the other hand. By achieving the objectives of such syllabi, not only can language teachers enhance the self-esteem and self-efficacy of language learners, but also such syllabi can help them take the proper path for their future progress in both academic and vocational settings.

Therefore, as one of the premises of a life syllabus, language learners will develop such abilities beyond the language skills that will stress their higher-order cognition and can empower them with more life qualities (Pishghadam & Zabihi, 2012). Such beyond-language skills are recommended for helping the learners to outperform outside as well as inside the classroom, may be of various subcategories. As one of those aforementioned life survival abilities, delving into the life difficulties and specially a critical modern issue such as 'depression' can bear the essential potentiality to be dealt with and overcome by the ongoing process of language education. Since language teaching is an integral part of the present-day education, finding ways by which language classrooms and activities become engaging enough for learners to get over life difficulties like depression seems highly rewarding.

One crucial function of life syllabus is according to Pishghadam and Zabihi (2012) to consider ELT as a "*life-and-language* enterprise, giving priority to the quality of learners' lives in advance of enhancing their language-related skills" (p.12), the language classrooms can be relocated to the places whereby the cognitive as well as pragmatic skills of language learners are promoted. Through applying life syllabus in language classrooms, a language learner can build a type of



unique identity for himself since according to Edwards (2009), the process of identity formation is a dynamic rather than static one and an individual' identity is liable to change during his social activities. Thus, the identity that a language learner builds in a language learning class can have as its basic constructing material the proficiency in English language as well as the attainment of life-related skills which was not accomplished by the previous syllabi to the present day.

Moreover, in this theory it is pointed out that teacher's role is about to change. In other words, this concept of education can mitigate the traditional role of teacher and lead him/her to be mostly emotionally engaging as well as to be reflective practitioners. Such practitioners can develop in training courses that enable teachers to become mostly educational teachers who would somehow or another 'go beyond teaching language per se towards extending their knowledge of other disciplines so that they can help learners develop as 'whole-person' individuals (Pishghadam, Zabihi & Kermanshahi, 2012, p. 893).

Alongside the importance of teacher in this theory, Pishghadam (2011) pinpoints the supremacy of language teaching classrooms over the other classrooms. Among numerous features that he offers, one highlights the opportunity such classes provide for learners in order to express themselves more freely. Thus as he concludes language learners have the privilege of assuming other identities that can move them forward so much so that they can turn to a less closed and less anxious person. This can be considered as a fresh start for them to get relieved of the previous life problems affecting both their personality disorders and learning problems.

### **3. Educational Therapy**

People with different personality disorders often seek help from the professional counselors in the field who can alleviate the pain and open new windows to a brighter future. As novel as promising guidelines such as those ones presented by Applied ELT and Life Syllabus have the potentiality to transform the learning context to that of a 'therapy' session in which the problematic learner is learning a new skill and getting prepared for outside the classroom life challenges concurrently. The notion of Educational therapy has been proposed by Caspari (1970) based on the principles of psychodynamic which integrates psychoanalytic therapy with teaching.

The great matter of importance here is the relationship which is built between a teacher as a therapist and a problematic learner. The areas of concern in educational therapy according to Caspari include difficulty in learning and communication, poor social behavior and building relationships, depression, deviation or other mental health problems (Jarvis, 2005). Also, most of the other therapy sessions are conducted in a heavy, unpleasant atmosphere which is time-limited,



expensive and is aimed at solving one aspect of a problem at a time, often place more pressure on the individual. The language teacher, on the other hand, can make a good investment by forming a productive relationship with the learners regardless of the aforementioned limitations.

### *3.1. Compensation hypothesis*

The therapist-patient relationship that is established for with the aim of helping a learner to overcome his/her affective problems may be hard to get outside of the classroom. This relationship takes time to be built by the necessary blocks of trust, in other words, the path by which a teacher can earn the trust of a learner may be much shorter than that of other formal therapists. According to Jarvis (2005), the incremental progress of the relationship between a teacher and a problematic learner can truly compensate for the lack of parental relationship which can be of vital importance to the learner. In such a fashion, the attachment that a teacher develops with a learner, if be of a secure type can compensate for any of those insecure ones the learner has with others such as his/her family.

### *3.2. Psychodynamic therapy*

Psychodynamic therapy is based on the assumption that a person is depressed because of unresolved, generally unconscious conflicts, often stemming from childhood. The goal of this type of therapy is for the patient to understand and cope better with these feelings by talking about the experiences. Psychodynamic therapy is administered over a period of weeks to months to years. Thus it seems that by considering the vital role of language classrooms proposed by Applied ELT, teachers can have the potentiality to prepare the language learners with some life related skills as well as building a therapeutic relationship with them.

Since some language learners seek to adopt a different identity from what they already have, especially those with depressive problems can get to develop a more productive relationship with their language teacher who while teaching the language-related skills, takes on a therapist role trying to peel the layers of learners' problems.

## **4. Discussion**

As was mentioned, Humanistic approaches in English Language Teaching (ELT) by placing much focus on the learner as a whole person and helping develop self-identity for him could direct the learner to that life-wise side of learning that was ignored by other approaches. All these



approaches, in a way or another, buttress the argument for a humanistic trend that is 'education for life' which is not well-applied in ELT.

We can lead to this conclusion that the previous syllabi in ELT, ranging from the 'product-oriented' ones focusing on structure, lexicon, or function-notion to the 'process-oriented-ones whose aim was to learn while accomplishing a task with the learner as the center of attention (Nunan, 1988), did not suffice to adhere to the philosophy of 'education for life' which would help the learners develop their life-wise qualities. Here, it seems well-grounded that having passed through different types of syllabi, ELT still needs a comprehensive and innovative kind of syllabus whose application can best define the notion of "ELT for life" (Pishghadam & Zabihi, 2012). Depression as one personality disorder can hinder and often paralyze a person in doing all his daily activities. Keeping depressed people in the centers which treat them as aliens and real '*patients*' can be more depressing in itself.

On the other hand, language classrooms as being rich in care and attention can provide suitable opportunities for troubled learners such as the ones with depression disorder. According to Educational therapy, building a friendly relationship with a depressed learner can lead to turn the classroom to one of a therapy sessions in which the focus is not only overcoming depression but also becoming able to communicate in another language which is less demanding and intimidating on the part of the learner. Such learning context, without imposing any onerous burden on the learner, can compensate for any other insecure contexts of living which can be more destructive than productive. Thus by following the principles of applied ELT and conducting a language class in which Life syllabus is the center of attention, a teacher with a wider scope can prepare and *treat* his learners for life-wise crises. The idea of using applied ELT, life syllabus in an educational therapy classroom can open new windows for handling the problems of all language learners, from the healthy to the troubled ones.

## References

- Beck, A.T., Rush, J., & Shaw, B.F. (1979). *Cognitive therapy of depression*. New York: Guilford Press.
- Carroll, D.W. (2008). *Psychology of language*. Belmont: Thomson Wadsworth.
- Dornyei, Z. (2005). *The psychology of the language learner*. Mahwah NJ: Lawrence Erlbaum Associates.
- Dornyei, Z. (2009). *The psychology of second language acquisition*. Oxford: Oxford University Press.
- Edwards, J. (2009). *Language and identity: An introduction*. Cambridge: Cambridge University Press.



- Jarvis, M. (2005). *The psychology of effective learning and teaching*. London, UK: Nelson Thrones Ltd.
- Kaplan, P.S. (1990). Educational psychology for tomorrow's teacher. St Paul, MN: West.
- Lipsey, M.W., Wilson, D.B. (1993). The efficacy of psychological , educational, and behavioural treatment: Confirmation from meta-analysis. *American psychologist*, 48(12), 1181-1209.
- Merrill, K.A., Tolbert, V.E., & Wade, W.A. (2003). Effectiveness of cognitive therapy in a community mental health center: A benchmarking study. *Journal of consulting and clinical psychology*, 71(9), 204-209.
- Nunan, D. (1988). *Syllabus design*. Oxford: Oxford University Press
- Oxford, R.L., & Ehrman, M.E. (1993). Second language research on individual differences. *Annual Review of Applied Linguistics*, 13, 188-205.
- Pishghadam, R. (2011). Introducing applied ELT as a new approach in second/foreign language studies. *Iranian EFL Journal*, 7(2), 8-14.
- Pishghadam, R., & Zabihi, R. (2012). Life syllabus: A new research agenda in English language teaching. *Perspective*, 1(19), 23-27.
- Pishghadam, R., Zabihi, R., & Kermanshahi, P. (2012). Educational language teaching: A new movement beyond reflective/critical teaching. *Life Science Journal*, 9(1), 892-899.
- Spolsky, B., & Hult, F.M. (2008). *The handbook of educational linguistics*. London: Blackwell.
- Taylor, G.R., & MacKenney, L. (2008). *Improving human learning in classroom*. Lanham: Rowman & Littlefield Education.
- Whitehead, A. N. (1929). *The aims of education and other essays*. New York, NY: Macmillan.
- Williams, M. & Burden, R. (2002). *Psychology for language teachers*. Cambridge University Press: Cambridge.

### Bio Data

**Azar Hosseini Fatemi** is associate professor in TEFL. She is the head of the Department of English, Ferdowsi University of Mashhad, Iran. She has supervised a large number of theses of MA and PhD graduates and published papers in ELT journals. Her areas of interest include issues in second language teaching and learning.

**Safoora Navari** is a PhD candidate in TEFL at Ferdowsi University of Mashhad. Having majored in English Language Literature for B.A. degree, she got her masters in TEFL both at the same university. Her areas of interest include teaching English and conducting research in the fields of psycholinguistics, sociology and discourse analysis. She is currently teaching English courses at Ferdowsi University and Payamnoor University of Mashhad.



## Does Differentiated Instruction Really Make a Big Difference?

Parviz Alavinia<sup>1</sup>

Sima Farhady

*Urmia University*

### Abstract

The current study sought to investigate the possible effects on vocabulary learning of the implementation of differentiated instruction (in the light of learners' multiple intelligences and learning styles). The research was carried out in the Iran Language Institute (ILI) in Urmia, with 80 Iranian intermediate female learners. Successive to homogenizing the groups, a total of 60 learners were found to enjoy the conditions for the rest of the research, and hence two 30-member groups were formed. These learners, then, sat the pretest (a vocabulary achievement test), and based on the results of multiple intelligences and learning styles questionnaires, which were administered later, the learners were divided into five separate categories termed visual-spatial (V), linguistic-auditory (L), kinesthetic-bodily (K), interpersonal (Inter), and intrapersonal (Intra). Next the learners in the experimental group were instructed based on their unique intelligence and learning style types, whereas the other class was taught in the traditional way with no differentiation. Subsequent to the administration of posttest the results were analyzed through the use of independent samples T-test and ANOVA. In line with the findings of the research, a significant amount of difference was found between the performances of two groups and in favor of the experimental group. Further, the performance of different learners with various intelligences and learning style was shown to vary significantly.

**Keywords:** Differentiated instruction, Learning styles, Mixed ability classes, Multiple intelligences, Vocabulary learning

### 1. Introduction

Though the belief that learners learn differently has now turned to a commonsensical view, notions such as one-size-fits-all instruction based on age-grade groupings, whole-class lecture teaching, and lockstep progress continue to be regarded as the norm even in today's revolutionized pedagogical world (Hess, 1999; Sizer, 1999). Nonetheless, as a glance through a typical classroom setting reveals, invisible diversities of learning characteristics and preferences dominate the ELT practices all over the globe. Some students come to school with little support and encouragements from home, while others commence the learning process, with skills and knowledge years beyond

---

<sup>1</sup> Email: [pevinia2006@yahoo.com](mailto:pevinia2006@yahoo.com)



grade level expectations (Tomlinson, 1999). As Caine and Caine (1990) note, "There can be up to a five-year difference in maturation between any two 'average' children" (p. 2).

Teachers and educators all over the globe have long reached consensus over the fact that learners in any given class are characterized by a wide variety of individual differences. Yet, the mere belief in these differences is not adequate and what the teachers need to do is to adjust their instructional and evaluative endeavors to these different learner needs and preferences. As Tomlinson (2000b, p. 1) contends, "Students who are the same age differ in their readiness to learn, their interests, their styles of learning, their experiences, and their life circumstances." Learner differences can be separated into two broad categories: demographic differences and individual learning differences. As a brief glance through the literature helps reveal, throughout the last two decades, many educational researchers have explored individual differences in learning, creating theories of thinking, learning, and teaching (Brooks & Brooks, 1999; Caine & Caine, 1990; Duckworth, 1996; Dunn, 2001; Guild, 1994; McCombs, 2001; Powell, 2000; Silver, Strong, & Perini, 1997; Sylwester, 1994).

Out of the manifold researchers' endeavors aimed at gauging the concept of learner differences, emerged the now-prevalent theory of differentiated instruction (Tomlinson, 1999, 2001, 2009a/b, 2010a/b), which intends to set forth a methodology for dealing with the so called learning differences (Hess, 1999). Thus, differentiated instruction was primarily given birth through the attempts by various researchers to address the diverse needs of learners in heterogeneous classroom settings (Westberg & Archambault, 1997). As Willis and Mann (2000, p.1) put it, "Differentiated instruction is a teaching philosophy based on the premise that teachers should adapt instruction to student differences.... teachers should modify their instruction to meet students' varying readiness levels, learning preferences, and interests." Also, as Stradling and Saunders (1993) maintain, differentiated instruction is "the process of matching learning targets, tasks, activities, resources, and learning support to individual learners' needs, styles, and rates of learning" (p. 129). Finally, to adopt Valiande, Kyriakides, and Koutselini's (2011) words,

Differentiation constitutes an innovating, constant reflective procedure of effective teaching and learning that cannot be met by readymade lesson plans. The planning and the instructional choices of a lesson plan based on differentiation can only be used effectively when chosen by the teacher, according to students' needs and other personal characteristics (p. 3).

Teaching vocabulary as one of the major categories of language learning is a broad area in which the learners' different learning characteristics play an important role. It goes without saying that



learners learn the vocabulary differently by resorting to the differential, personalized styles and strategies they possess; thus, to be able to implement a practical method of teaching vocabulary, teachers need to be aware of their learners' current skill levels, strengths and challenges, interests and preferences, and needs and goals. The challenge is for teachers to ensure that the needs of all learners are equally valued and attended to.

The purpose of this study is to see if using differentiated instruction would have a positive effect on students' vocabulary learning in mixed ability classes. When first starting to design differentiated instruction, it is important that all teachers and support staff be willing to devote the time, energy, and resources to design and implement a successful program. Roles need to be clearly defined, schedules need to be made in advance and adhered to as much as possible, and collaboration/cooperation needs to replace competition at all levels. In the beginning, redesigning curriculum and teacher roles for differentiated instruction is very intensive labor. However, using technology, future planning can be streamlined, teachers can use what was successful in the past and adjust activities that may need more clarity, and spend more time on the instruction of each student according to her/his needs.

Being involved with the teaching career as colleagues since long, the researchers had been pondering over the idea of different learning styles for years, especially as facing those learners whose learning styles are so clear that can be readily recognized by the instructor. Perhaps the first trigger for the current study came from their sensitivity toward the age-old premise that a typical class is really made up of learners who enjoy a range of varied proficiency levels, and come from different backgrounds.

Today's classrooms are becoming more academically diverse. Many, if not most, classrooms contain students representing both gender and multiple cultures and generally contain students with a range of exceptionalities and markedly different experiential backgrounds. These students almost certainly work at differing readiness levels, have varying interests, and learn in a variety of ways. In order to create a kind of educational space which is beneficial for all students with different intelligences and styles, the teaching methods are to be changed to fit all the students.

Though the domains of learning styles and multiple intelligences have long made their ways into the field of applied linguistics and an assortment of varied research projects have been targeted toward these partially novel areas of investigation, it seems that scant attention has been paid to the possible effect of differentiated instruction via paying due heed to these two variables on learning vocabulary. Thus, the current study intends to investigate this partly untouched area in a brand new context, i.e. for Iranian intermediate EFL learners.



In tandem with the principal axioms of the current investigation, the following questions were set forth:

1. Does differentiated instruction have any impact on Iranian EFL learners' vocabulary learning?
2. Is there any significant difference between the performance of learners with different learning styles (auditory, visual, and kinesthetic) on the vocabulary posttest?
3. Is there any significant difference between the performance of learners with different multiple intelligences (spatial, linguistic, bodily-kinesthetic, interpersonal, and intrapersonal) on the vocabulary posttest?

## **2. Theoretical Background**

Diversification and metamorphosis in today's revolutionized educational system is an inevitable verity, which has partly emanated from the avant-garde socio-cultural requirements of ultramodern societies. In the current state of affairs, we are no more faced with the conventional sameness (in terms of learners' abilities, preferences, needs, traits, and the like), which once used to dominate the mainstream educational arenas. Indeed, our present-day didactic endeavors are characterized by intricacies germane to intragroup diversities, in the sense that individuals with whom we are currently dealing are no more the uni-dimensional beings of, say, 1800s, as Tomlinson (1999) contends; rather than resemblances which constituted the commonsensical norm in primordial pedagogical enterprise, it is the notion of disparities that has been and is to be given more heed in the modern practice of teaching and learning.

Nevertheless, the main challenge for current pedagogy is not the mere confession to the prevalence of differences in the learning context. Awareness of such variations and endorsement of learner differences can just feature as the springboard for a more critical stage, which is the implementation of sound approaches and methods aimed at addressing the complicated needs of modern pedagogy. Though a plethora of attempts have, thus far, been set in place by a multitude of educationalists, practitioners, and theoreticians, a great deal still remains to be done toward the diversification of instructional/evaluative materials and techniques, in keeping with the demands of today's diversified learning environments.

In the campaign targeted toward the reappraisal of teaching/learning practices with the aim of meeting individual learning differences, later on being referred to by the neologism differentiated instruction, perhaps it was Gardner (1983, 1991, 1993) who initially brought the notion of instructional diversification to the foreground of attention via the establishment of his groundbreaking theory of multiple intelligences. Another line of theories concerning the notion of



differentiation in instruction seems to have been configured out of the attempt by several prominent researchers in the field of learning styles (e.g. Chislett & Chapman, 2005; Dunn, 2000, 2001; Dunn & Dunn, 1992a/b; Honey & Mumford, 2006; Kolb, 1985).

Literature is replete with the research projects reporting on the positive outcomes resulting from the use of differentiated instruction. Yet, owing to the partial novelty of the concept of differentiation in instruction, more time might be required on the part of (novice) teachers and learners to come into terms with the underlying tenets of this groundbreaking theory. An early case study of one middle school's experience with differentiated instruction by Tomlinson (1995), for instance, came up with the conclusion that there was a need to investigate teacher resistance to new models catering for academic diversity, as well as considering teachers' perception of classroom management in the light of these changes.

It is, therefore, mostly the case that teachers who experience early successes with differentiation are more prone to adaptation in this respect. In a study investigating the impact of differentiated instruction on student scores on standardized tests, Hodge (1997) found that students who were prepared for tests using differentiated techniques showed a gain in their mathematics scores, but there were no comparable gains in reading scores. However, teachers' perceptions of being able to meet the needs of diverse learners, as reported in this investigation, did not appear to be affected by the use of either traditional or differentiated instructional techniques.

In another study, Tomlinson, Moon and Callahan (1998) investigated the nature of instructional practice among middle school populations, considering the degree to which teachers respond appropriately to academic diversity, using differentiation. This study indicated that very few teachers take students' interests, learning profile or cultural differences into account when they plan lessons. As the findings of this study revealed, though some of the teachers who used varied instructional strategies partly succeeded in creating more flexible classroom settings, most teachers expressed frustration about attempting to deal with learner variance.

A word of warning is to be made, at this juncture, regarding some teachers' misconception regarding the real practice of differentiated instruction, as it is sometimes the case that, on the face of it, the teacher claims to be dealing with differentiation, whereas in actuality s/he does nothing more than tracking the traditional approach. This is evident in a study by Blozowich (2001) who found that teachers used a variety of techniques but continued to prepare lessons as they would for a tracked classroom. This researcher, then, came to the conclusion that teachers implementing differentiated instruction are in dire need of incessant and consistent professional development, along with intensive discussion about how these techniques are to be put to use in the classroom.



Among the other supportive claims in favor of the privileges of differentiated instruction, reference can be made to McAdamis' (2001) work, which pointed to significant performance enhancement among the students in the Rockwood School District (Missouri), successive to the implementation of differentiated instruction. Planning, mentoring, professional development, doing action research and holding workshops are highlighted in this study as the principal measures to be taken with the aim of lessening teachers' initial resistance toward differentiation in education.

In another study concerning the repercussions of adopting differentiated syllabi for teachers, Affholder (2003) found that teachers who made an intensive use of these strategies developed a better individual perception and assumed greater responsibility for student growth. Furthermore, in line with the results of this study, teachers who were characterized by a more frequent utilization of differentiated techniques in their classes came up with higher levels of self-efficacy as well as willingness to try new instructional approaches. Yet, as the results revealed, those enjoying seniority and higher experience were found to be at a more privileged position in this respect. Working with the unique community of undergraduate teachers, Johnsen (2003) concluded that the use of differentiated techniques proved to be a beneficial tool in keeping the individuals' interested and, consequently, the implemented techniques were said to have provided the undergraduate teachers with a highly gratifying experience.

Despite the great body of work carried out on varied gains resulting from the implementation of differentiated instruction, Robison (2004) calls for further research into the utilization of differentiated instruction techniques, mainly owing to the fact that teachers are still reluctant to pursue the axioms of differentiation, due to an assortment of reasons like the problem with setting time constraints for the productive use of such techniques in classes.

Nevertheless, evidence for mounting (renewed) interest in and attention toward the importance of differentiated instruction might be gathered via a fleeting glimpse through the recent probes allotted to this area of research, among which mention can be made of McQuarrie, and McRae's (2010) project which reports on the outcomes of a Western Canadian research review regarding the effective practices supporting differentiated instruction. Based on the findings of the study, these researchers claim that "addressing student diversity and providing the best learning opportunities for all children across Kindergarten to Grade 12 schools requires recognition that differentiation requires time, training, intentional planning and long-term commitment on the part of educators, government and wider school communities" (McQuarrie & McRae, 2010, p. 1)

Finally, in a more recent investigation aimed at the practical implementation of differentiated instruction, Valiande, Kyriakides, and Koutselini (2011) organized a probe into "a) the effects of



differentiated instruction on students achievement in mixed ability classrooms and b) the dimensions of quality and equity effectiveness of differentiated instruction, implemented by Cypriot teachers in order to improve their effectiveness" (p. 6). As they continue to aver, in line with the gained upshots, they were successful in providing strong "evidence to support the theory of differentiation, its basic principles and presuppositions, and confirm its effectiveness in mixed ability classrooms" (Valiande, Kyriakides, & Koutselini, 2011, p. 13).

### **3. Method**

#### *3.1. Participants*

Sixty Iranian female learners studying at the ILI (Iran Language Institute) in Urmia, participated in this study. They were all intermediate students and their age ranged between 15 and 20. Actually these sixty participants were in two separate classes, with each class including thirty participants. In order to group them according to their type of intelligences and learning styles, a multiple intelligences test and a learning styles test were initially undertaken.

Though, based on the institute's placement procedures, all the participants were supposed to enjoy roughly the same level of proficiency, to ascertain the homogeneity of the group, a standardized test entitled Preliminary English Test (PET) was used to ascertain the homogeneity of participants at the outset of the study. The participants were selected out of a pool of 80 learners according to their performance on the PET. Through considering the normal distribution of the participants' scores on proficiency test, merely those whose scores were one standard deviation above and one standard deviation below the mean ( $M = 40.56$ ) of the normal distribution curve were chosen for the study. Consequently, 60 of them were found to be homogenous and were chosen for the purpose of the current study.

#### *3.2. Instrumentation*

To commence the study, the Preliminary English Test (PET), a second level Cambridge ESOL exam for intermediate level learners, was administered to ensure the homogeneity of the subjects in terms of language proficiency. The test included four sections of Reading, Writing, Listening, and Speaking. Those whose scores fell one standard deviation above and below the mean (i.e., between 44.40 and 36.72 out of 50) participated in the study.

Furthermore, to obtain the objectives of this study, the participants (in both experimental and control groups) were given an intermediate vocabulary test (used both as pre-test and post-test), which was



adopted from the test book designed particularly for this level at the institute where they studied. The test consisted of 50 multiple choice questions and was utilized to test whether differentiated instruction had any impact on the participants' vocabulary learning or not.

Moreover, to determine the participants' types of intelligences, a multiple intelligences test (based on Howard Gardner's MI Model) was given to them. The test consisted of 34 questions which were scored on the scale of 1 to 4.

VAK Learning Styles self-assessment questionnaire was the other test given to the participants in order to determine their learning styles. The questionnaire consisted of 30 multiple choice questions indicating visual, auditory and kinesthetic learning styles.

### *3.3. Design*

As the present research involved selecting one group upon which a variable was tested without any random pre-selection processes and included a treatment on the experimental group, the design was quasi-experimental and included pretest/posttest to evaluate whether there were significant differences in the vocabulary achievement of the participants.

As stated previously, at the outset of the study, the Preliminary English Test (PET) was administered to learners to ascertain the homogeneity of participants in terms of language proficiency. Then, the selected participants were given the pretest before the application of any differentiated instruction and the obtained scores in both classes were collected for later analysis. In order to apply differentiation, the determination of the participants' intelligences and learning styles types was required. Therefore, the multiple intelligences test and the learning styles questionnaire were given to the participants. In both classes the participants were grouped based on their type of intelligences and learning styles.

Next, one of the classes was randomly selected to undertake differentiated instruction in which different groups of participants enjoying the same kind of intelligences and learning styles were instructed according to their special type of intelligences and learning styles for about seventeen sessions. As a consequence of this kind of grouping applied for the experimental group participants, five distinct groups were formed (with each group encompassing 6 participants). These subgroups were termed visual-spatial, linguistic (auditory), bodily-kinesthetic, intrapersonal and interpersonal, in accordance with the intelligences and learning styles they possessed.

Thus, each group was instructed based on the mentioned types of intelligence and learning style for a matter of about 17 sessions. Each session lasted for about 90 minutes. The other class was just



instructed in the traditional way without any differentiation. At the end of the course a post test encompassing 50 multiple-choice vocabulary items was given to the participants to test whether the use of the differentiated instructions had a positive impact on the participants' learning or not. The obtained scores from the two tests (pre-test and post-test) were collected for the data analysis section in the study. The researcher assigned the learners' scores into 5 different groups (as mentioned above) based on their type of intelligence and learning style.

### *3.4. Data collection and analysis*

The current study set out with the aim of gauging the possible impact of the application of differentiated instruction on Iranian EFL learners' vocabulary learning, with a focus on learners' multiple intelligences as well as their learning styles. To this end, four types of data were collected: 1) the data showing the results of the Pet Test to homogenize the participants, 2) the data indicating the participants' unique intelligences and learning styles, 3) the participants' scores on the two tests (pretest and posttest), and 4) the data showing the performance of each group of learning styles and multiple intelligences on the tests.

As stated earlier, the same 50-item test was utilized as both pretest and post-test. Each question carried 2 points, and hence the total score of the test was 100. The control participants' scores on the pretest ranged between 60 and 90, and on the posttest, their scores were between 70 and 90. Yet, the experimental participants' scores on the pretest were between 60 and 95, and on the posttest, their scores ranged between 80 and 100. The data obtained were mainly processed and analyzed through the use of t-test and ANOVA.

## **4. Results**

### *4.1. Results gained for research question one*

1. Does differentiated instruction have any impact on Iranian EFL learners' vocabulary learning?

Table 1 represents the information relevant to the performance of control group. As the table reveals, the mean score of the participants who took the pretest in the control group (PRT1) equaled 79.5, and their mean score on the posttest (POT1) was 82.5.



**Table 1.**

Descriptive statistics for pretest and posttest in the control group

GROUP	N	Mean	Mean difference
PRT1	30	79.5	-3.0
POT1	30	82.5	

Table 2 shows the findings obtained regarding the performance of experimental group. As is seen, the mean score of the participants who took the pretest in the experimental group (PRT2) was 81.5, and their mean score on the posttest (POT2) equaled 89.3.

**Table 2.**

Descriptive statistics for pretest and posttest in the experimental class

GROUP	N	Mean	Mean difference
PRT2	30	81.5	-7.8
POT2	30	89.3	

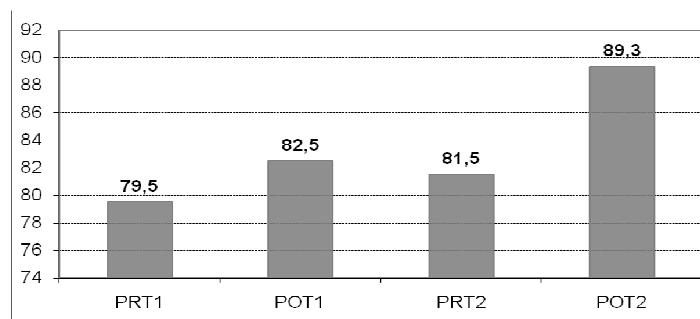
The comparison between the two mean scores of each class can be utilized as an insight toward the rejection of the first null hypothesis claiming no difference between the performances of two groups as a result of the treatment applied. The slight difference between the mean scores of the PRT1 and POT1 in comparison with the noticeable difference between the mean scores of the PRT2 and POT2 can be taken as evidence in favor of the claim that the application of the differentiated instruction in the second class (the experimental group) has been effective. Yet, to make a stronger claim, more robust evidence is called for; as the data presented in Table 3 help reveal ( $df = 29$ ,  $p = 0.00$ ), a significant amount of difference is witnessed with regard to the performance of two groups. Thus, it can be concluded that differentiated instruction has had a positive effect on the students' vocabulary learning and the first hypothesis is, therefore, prone to rejection. Moreover, Figure 1 helps shed further light on the claim that the difference between the mean scores of POT1 and POT2 is indicative of the noticeable change in the participants' performance in the experimental class.



**Table 3.**

Independent samples t-test for experimental group in PRT2 and POT2

	t	Df	Sig (2-tailed)	Mean Difference	95 % confidence interval of Difference	
					Lower	Upper
Control	54.65	29	0.00	81.55	73.541	83.64
Experimental	40.51	29	0.00	85.50	74.231	99.12



**Figure 1.**

The investigation of the mean scores in two classes

#### 4.2. Results gained for research question two

2. Is there any significant difference between the performance of learners with different learning styles (auditory, visual, and kinesthetic) on the vocabulary posttest?

As stated earlier, to come up with a firm basis for performing the groupings within the experimental group, use was made of the results gained through the implementation of both learning styles inventory and multiple intelligences test. Altogether, five categories of learners were constituted drawing on the test results, and were alternatively termed: visual-spatial (V), linguistic-auditory (L), kinesthetic-bodily (K), interpersonal (Inter), and intrapersonal (Intra). Yet, the more convenient type of grouping for the sake of data analysis was thought to be the one which assigned learners to different categories, based on their learning styles or multiple intelligences separately.

Table 4 depicts the descriptive statistics gained for the performance of learners with different learning styles. As can be inferred from the table, the mean of visual learning style scores on the



posttest of vocabulary achievement test is the highest among the three types of learning styles; yet its Std. deviation is the lowest in comparison with the other two groups.

Also, drawing on the findings presented in Table 5 ( $df = 5$ ,  $p = 0.00$ ), it can be concluded that there is a significant amount of difference among the performance of learners possessing different learning styles (auditory, kinesthetic and visual), and hence the second hypothesis claiming no difference between the three groups of learning styles can be disconfirmed.

**Table 4.**

Descriptive statistics relevant to the performance of three learning styles

Group	N	Mean	Std. Deviation	Std. Error Mean
A	6	86.2500	14.12002	5.76447
V	6	89.1667	5.84523	2.38630
K	6	82.9167	7.14435	2.91667

**Table 5.**

ANOVA results for different learning styles

Group	t	df	Sig (2-tailed)	Mean Difference	95 % confidence interval of Difference	
					Lower	Upper
A	14.962	5	.000	86.2500	71.4319	101.0681
K	37.366	5	.000	89.1667	83.0325	95.3009
V	28.429	5	.000	82.9167	75.4191	90.4142

#### *4.3. Results gained for research question three*

3. Is there any significant difference between the performance of learners with different multiple intelligences (spatial, linguistic, bodily-kinesthetic, interpersonal, and intrapersonal) on the vocabulary posttest?

To analyze the third research question of the study, the main grouping of the study with five subcategories was employed, but with a number of different labels. Hence, the five categories on which the analysis was carried out were spatial (S), linguistic (L), bodily-kinesthetic (K),



interpersonal (inter), and intrapersonal (intra). As the data in Table 6 help reveal, among the five groups of intelligences, the highest and lowest posttest means belonged to spatial ( $M = 89.1667$ ) and interpersonal ( $M = 80.421$ ) intelligences, respectively; yet, the greatest and lowest amounts of standard deviation were gained for linguistic ( $SD = 14.12002$ ) and spatial ( $SD = 5.84523$ ) types of intelligence, respectively.

As it is shown in Table 7, ( $df = 5, p = .000$ ), the difference among the performance of five different groups of multiple intelligences is statistically significant, and, therefore, the third null hypothesis postulating the non-existence of difference among the performance of various intelligence types on the vocabulary posttest can be nullified. Figure 2 gives out a more lucid view of this significant difference, showing that spatial intelligence group had the best performance in terms of vocabulary achievement test.

**Table 6.**

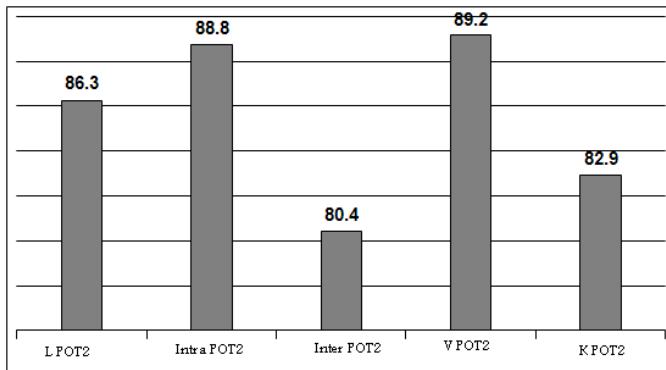
Descriptive statistics relevant to the performance of five intelligence types

Group	N	Mean	Std. Deviation	Std. Error Mean
L	6	86.2500	14.12002	5.76447
S	6	89.1667	5.84523	2.38630
K	6	82.9167	7.14435	2.91667
Intra	6	88.801	6.102327	2.51092
Inter	6	80.421	10.80123	6.79021

**Table 7.**

ANOVA results for different intelligence types

Group	t	df	Sig (2-tailed)	Mean Difference	95 % confidence interval of Difference	
					Lower	Upper
L	14.962	5	.000	86.2500	71.4319	101.0681
S	37.366	5	.000	89.1667	83.0325	95.3009
K	28.429	5	.000	82.9167	75.4191	90.4142
Intra	34.574	5	.000	88.801	76.0520	92.4582
Inter	26.601	5	.000	80.421	73.812	90.1270



**Figure 2.**

The investigation of the mean scores of the five different groups of multiple intelligences

## 5. Discussion

The main objective of this study was to investigate the effect of Differentiated Instruction on the students' vocabulary learning. According to the results of the study relevant to the first research question, the researchers concluded that the application of differentiation had a positive impact on the learners' performance, and these findings were in line with some previous research done by Affholder (2003), Gartin, Murdick, Imbeau, and Perner, (2002), Hodge (1997), Johnsen (2003), and McAdamis (2001).

Gartin et al. (2002) conducted a study in which differentiated instruction was a method for developing curriculum for successful learning and instruction and supported the application of the differentiated instruction. Johnsen (2003) carried out an investigation using differentiated instruction with undergraduate teachers to suit different ability levels. Student teachers in this context were encouraged to differentiate content and process, using learning centers, different reading materials and different strategies. Based on the gained upshots, the implemented techniques were found to have provided the undergraduate teachers with a highly gratifying experience

McAdamis (2001), on the other hand, reported significant improvement in the test scores of low-scoring students in the Rockwood School District (Missouri), following the use of differentiated instruction. Hodge's (1997) study revealed that students who were prepared for tests using differentiated techniques showed a gain in their mathematics scores. Finally, Affholder (2003) found that teachers who used differentiated instruction strategies more intensively showed improved individual perception and adopted greater responsibility for student growth.



As to the second and third research questions, which sought to investigate the possible differences among the performance of different groups of learning styles and multiple intelligences, and substantiated the significant difference in this respect, no support or counterevidence was encountered within the previous body of literature, mainly owing to the fact that the researchers were not able to find any relevant research addressing the issue of learner differences compliant with the perspective taken in the current study.

All in all, the current study applied Differentiated Instruction to teaching vocabulary in mixed ability classes with a focus on the learners' learning styles and multiple intelligences, and, by way of doing so, investigated the effect of differentiated instruction on learners' vocabulary learning. As far as the difference between the two kinds of instruction, i.e. differentiated and traditional, are concerned, the results indicated that the class in which differentiated instruction was conducted, had a significantly better performance on the post-test and consequently it was concluded that the application of differentiated instruction had a positive effect on the process of learning. Furthermore, with regard to the second and third research questions, which probed the viable differences among different groups of learners possessing different learning styles and multiple intelligences, significant differences were found to be at work. Finally, another intriguing piece of finding gained in the current study was the observation that participants in the experimental group felt more comfortable and confident with regard to their capability for achievement; in other words, differentiated instruction and having students work in small groups really helped them become more empowered learners.

## References

- Affholder, L. P. (2003). *Differentiated instruction in inclusive elementary classrooms*. University of Kansas, Kansas.
- Blozowich, D. G. (2001). *Differentiated instruction in heterogeneously grouped sixth grade classrooms*, doctoral dissertation, Immaculate College.
- Brooks, M., & Brooks, J. (1999). The courage to be constructivist. *Educational Leadership*, 57(3), 18-24.
- Caine, R., & Caine, G. (1990). Understanding a brain-based approach to learning and teaching. *Educational Leadership*, 47(2), 66-70.
- Chislett, V., & Chapman, A. (2005). *VAK learning styles self-assessment questionnaire*. Retrieved from <http://www.scribd.com/doc/4638815/vak-learning-styles-questionnaire>
- Duckworth, E. (1996). *The having of wonderful ideas and other essays on teaching and learning*. New York: Teachers College Press.
- Dunn, R. (2000). Learning styles: Theory, research, and practice. *National Forum of Applied Educational Research Journal*, 13(1), 3-22.



- Dunn, R. (2001). Learning style differences of nonconforming middle-school students. *NASSP Bulletin*, 85, 68-74.
- Dunn, R., & Dunn, K. (1992a). *Teaching elementary student through their individual learning styles*. Boston: Allyn & Bacon.
- Dunn, R., & Dunn, K. (1992b). *Teaching secondary students through their individual learning styles*. Boston: Allyn & Bacon.
- Gardner, H. (1983). *Frames of mind*. New York: Basic Books.
- Gardner, H. (1991). *The unschooled mind: How children think and how schools should teach*. New York: Basic Books.
- Gardner, H. (1993). *Multiple intelligences: The theory in practice*. New York: Basic Books.
- Gartin, B., Murdick, N., Imbeau, M., & Perner, D. (2002). *How to use differentiated instruction with students with developmental disabilities in the general education classroom*. Arlington, VA: Council for Exceptional Children.
- Guild, P. (1994). The culture/learning style connection. *Educational Leadership*, 51(8), 16-21.
- Hess, M. A. (1999). *Teaching in mixed-ability classrooms*. Retrieved from <http://www.weac.org/kids/1998-99/march99/differ.htm>
- Hodge, P. H. (1997). *An analysis of the impact of a prescribed staff development program in differentiated instruction on student achievement and the attitudes of teachers and parents toward that instruction*, Published thesis. University of Alabama.
- Honey, P., Mumford, A. (2006). *Learning styles questionnaire: 80 item version*. Peter Honey Publications: London.
- Johnsen, S. (2003). Adapting instruction with heterogeneous groups. *Gifted Child Today*, 26(3), 5-6.
- Kolb, D. (1985). *Learning style inventory*. Boston, MA: McBer and Company.
- McAdamis, S. (2001). Teachers tailor their instruction to meet a variety of student needs. *Journal of Staff Development*, 22(2), 1-5.
- McCombs, B. (2001). *Understanding the keys to motivation to learn*. Retrieved from <http://www.mcrel.org/products/noteworthy/noteworthy/barbaram.asp>
- McQuarrie, L. M., & McRae, P. (2010). A provincial perspective on differentiated instruction: The Alberta Initiative for School Improvement (AISI). *Journal of Applied Research on Learning*, 3(4), 1-18.
- Powell, M. J. (2000). How can research on the brain inform education? *Classroom Compass*, 3, Retrieved from <http://www.sedl.org/scimath/compass/v03n02/1.html>
- Robison, E. M. (2004). *Teacher decision-making in utilising differentiated instruction*. Unpublished PhD thesis. Marywood University.
- Silver, H., Strong, R., & Perini, M. (1997). Integrating learning styles and multiple intelligences. *Educational Leadership*, 55(1), 22-29.
- Sizer, T. (1999). No two are quite alike. *Educational Leadership*, 57(1), 6-11.



- Stradling, B. & Saunders, L. (1993). Differentiation in practice: Responding to the needs of all pupils. *Educational Research*, 35(2), 127-137.
- Sylwester, R. (1994). How emotions affect learning. *Educational Leadership*, 52(2), 60-66.
- Tomlinson, C. (1995). Deciding to differentiate instruction in middle school: One school's journey. *Gifted Child Quarterly*, 39(2), 77-87.
- Tomlinson, C. A. (1999). *The differentiated classroom: Responding to the needs of all learners*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Tomlinson, C. A. (2000b). Reconcilable differences? Standards-based teaching and differentiation. *Educational Leadership*, 58(1), 1-7.
- Tomlinson, C. A. (2001). *How to differentiate instruction in mixed-ability classrooms*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Tomlinson, C. (2009a). Intersections between differentiation and literacy instruction: Shared principles worth sharing. *New England Reading Association Journal*, 45(1), 28-33.
- Tomlinson, C. (2009b). Myth #8: The patch-on approach to programming is effective. *Gifted Child Quarterly*, 53, 254-256.
- Tomlinson, C. (2010a). One kid at a time. *Educational Leadership*, 67(5), 12-16.
- Tomlinson, C. (2010b). Differentiating instruction for academic diversity. In J.M. Cooper (Ed.). *Classroom teaching skills* (pp. 153-187). Wadsworth, Centgage Learning.
- Tomlinson, C. A., Moon, T. R., and Callahan, C. M. (1998). How well are we addressing academic diversity in the middle school? *Middle School Journal*, 29(3), 3-11.
- Valiande, A. S., Kyriakides, L., Koutselini, M. (2011). *Investigating the impact of differentiated instruction in mixed ability classrooms: Its impact on the quality and equity dimensions of education effectiveness*. Paper presented at the International Congress for School Effectiveness and Improvement. Springwood, New South Wales Australia.
- Westberg, K., & Archamabult, F. (1997). A multi-site case study of successful classroom practices for high ability students. *The Gifted Child Quarterly*, 41(1), 42-51.
- Willis, S., & Mann, L. (2000). Differentiating instruction: Finding manageable ways to meet individual needs (excerpt). *Curriculum Update*, ASCD. Retrieved from <http://www.ascd.org/readingroom/cupdate/2000/1win.html106>

## Bio Data

**Parviz Alavinia** holds a PhD in TEFL/TESOL from Allameh Tabatabai University. He is currently involved as a fulltime assistant professor and staff member at Urmia University. His main areas of interests include psycholinguistics, philosophy of language, (critical) discourse analysis, task-based instruction, and particularly emotional intelligence.



**Sima Farhady** holds an MA from Islamic Azad University (Urmia Science and Research Branch). She is mainly interested in the domain of differentiated instruction and individualization of education.



# The Effect of an Output-Packet Practice Tasks on the Proceduralization of Knowledge in Intermediate Iranian Foreign Language Learners

Mehroosh Fakharzadeh

Manidjeh Youhanaee<sup>1</sup>

*University of Isfahan*

## Abstract

This study investigates the effect of output practice tasks on the acquisition of English modals. The acquisition is measured in terms of evaluating the comprehension and production proceduralized skills. A pretest posttest and delayed posttest design was used where the procedural knowledge was specifically operationalized through the groups' performance on timed GJ and completion tests. Two intact classes of intermediate EFL learners were randomly assigned to treatment and control groups. The treatment group, here the output group ( $n= 27$ ) received a) explicit grammar instruction, b) a combination of three output tasks including dictogloss, individual text reconstruction, and corrected-close translation , and c) feedback. The control group ( $n=25$ ) were just exposed to the identical texts trough listening and reading tasks followed by some questions irrelevant to the target structure. Results showed that on the posttest, administered 10 days after the last treatment session, the output group outperformed the control group in both measures of procedural knowledge. Forty days later, the output group's performance was still significantly better than that on the pretest. This group also succeeded in retaining its outperformance on both measures of procedural knowledge delayed posttest. The results may help language teachers design more effective activities for the learners considering the institutional constraints.

**Keywords:** Declarative knowledge, Proceduralization, Practice, Transfer of knowledge

## 1. Introduction

The goal of teaching in real-world setting is to support two aspects of post-instruction performance: a) the relatively long-time retention of knowledge and b) the ability to generalize the instructed items to other contexts (DeKeyser, 2007; Schmidt & Bjork, 1992). However, while the goals may be relatively easy to agree on, the question of means is far more complicated. To address the optimal

---

<sup>1</sup> Email: youhanaee\_m@hotmail.com



ways to achieve the goals of instruction in general and those in language teaching in particular, researchers have focused on various approaches and aspects of instruction; some have emphasized developing instructional method and materials, others have discussed about the effect of setting, and yet a number of them have dealt with the amount and type of practice.

One currently popular approach to learning, skill acquisition, which considers a key role for practice, is skill acquisition theory. According to skill-building theory of learning, also known as ACT (Adaptive Control of Thought) family of skill acquisition, it is explicit knowledge that gives rise to implicit knowledge, when the former is proceduralized through practice. That is, learners first learn a rule as a declarative fact explicitly and then by practicing the use of this rule they convert it into an implicit representation (Anderson, 1983, 1993).

Anderson's ACT model, with its goal to provide explanation for various aspects of human cognition has been updated until last few years (Anderson, Fincham & Douglass, 1997; Anderson & Shunn, 2000). According to the last version of this model, ACT-R, following systems is the basic architecture of cognition: the perceptual-motor system, the goal system, the declarative memory and finally the procedural memory. The components which have been received lots of attention from educational psychologists have been the declarative and the procedural memory, since these two modules are involved when the function of practice is the focal point (Seranno, 2007).

Declarative knowledge is added into the memory through chunks and they become activated and subsequently acquire strength as a result of frequent use. What characterizes the later versions of this model is the role of examples. Not only declarative knowledge but also examples can be used as a source for procedural knowledge through the processes of analogy and knowledge compilation (Anderson et al., 1994). Proceduralization takes place when declarative knowledge is transformed into procedures, production rules, for performing a skill. Production rules also require practice to acquire activation and strength. Practice at this stage results in automatization of skill (Anderson, 1993).

This corresponds closely with the cycle 'presentation-practice-production', which is not unfamiliar to those who have been involved in language teaching (Byrne, 1976). In spite of the fact that most learning models have something equivalent to these three stages, they fell into disfavor for some time. At least, in the domain of second language acquisition, due to the popularity of communicative language teaching (CLT) and its emphasis on meaningfulness of class activities, language functions, and fluency in second language classes, practice and explicit knowledge have been overlooked by the second language teachers and researchers for decades (DeKeyser, 2007; Dornyei, 2009). The idea, however, faced an objection from the language teachers and researchers



warning that proficiency in a second language requires fluency as well as complexity and accuracy (Ellis, 2005; Skehan, 1998). This, together with issues such as implicit/explicit language learning (Bialystok, 1994) led into the resurgence of instructed second language acquisition and form-focused approach to teaching language.

The implications of the ACT family of skill acquisition can be employed in language teaching. In applied linguistics this was first formally advanced by Sharwood Smith (1981) and was subsequently promoted by DeKeyser (1998, 2007). Anderson (1993) suggests that to teach a skill one should start with declarative instruction, and examples. By analogy to those examples, the learners will gradually apply the production rules repeatedly until such behavior is strengthened and reliable. The instructor should monitor the learners' performance and provide feedback to make sure that they have acquired the right rule and its application.

Another issue which surrounds the duality of knowledge in ACT models is skill-specificity. The skill-specificity issue is probably the one that has drawn the most attention in applied linguistics lately (DeKeyser, Salaberry, Robinson, & Harrington, 2002; Izumi, 2002, 2003; Muranoi, 2007; Qin, 2008). It refers to the effect of practice on the proceduralization of declarative knowledge. The practice effect is said to be specific since after the initial practice of declarative knowledge, learners incorporate the knowledge into behavioral routines which are very specific rules and can be used fast with low rate of error (DeKeyser, 2007). The production rules that were created at the second phase of skill acquisition keep fine-tuning so that they become inflexible. It is believed that even a slight change in the condition or parameters of practice task might result in failure in performing on post-training tests (Schmidt et al., 1992). However, skill-specificity is most probably the result of years and years of practice and prior to automatization of skill there is no way to guarantee that it is speeded availability of learners' declarative knowledge resources or production rules which have been created and fine-tuned to that specific practice task. Yet, the extent to which the acquired knowledge can be generalized, as a result of instruction and practice, to other contexts, seems to be of high importance. This phenomenon, known as "Transfer of Learning" refers to how previous leaning influences current and future learning or performance, and how past or current leaning is applied to similar novel situations (Haskel, 2001, Speelman & Kirsner, 2005).

Despite the general assumption that people have different degrees of aptitude to perform specific skills, the influence of practice has been emphasized by some authors. Eriksson, Krampe and Tesch-Romer (1993), for instance, claim that it is years of effortful, deliberate practice, that distinguishes experts from novices. Yet, in a foreign language context, like Iran, language learners usually do not have the opportunity to be immersed for a long time in a rich target language environment. Rather, they have limited language input. That is why parallel to cognitive and



educational psychologists second language acquisition researchers are looking for optimal ways to deal with the insufficient time and opportunities to introduce a large amount of practice to learners, and have made attempts to manipulate one or more variables in relation to practice tasks, such as scheduling of practice, its type and last but surely not least its variety.

Manipulating the scheduling of practice, for instance, Bird (2010) compared the effect of distributed with massed practices on the learning of English simple past and past perfect tenses. For distributed practice the learners received practice at long intervals and for massed practice they received short ones. The study used isolated out-of-context form-focused materials as the material for practice. The results of the study revealed that the distributed practice resulted in superior long-term retention test of comprehension. Transferability of knowledge to novel context was not addressed in the study. Additionally, the measuring instruments of learning were form-focused requiring conscious attention of the participants. Therefore, the results cannot indicate proceduralization or automatization of declarative knowledge as a result of practice.

Serrano and Munoz (2007) investigated the effect of the distribution of instructional time on the acquisition of a second or foreign language. They analyzed the performance of adult students enrolled in three different types of EFL programs, 'extensive', which distributed a total of 110 h in 7 months, 'semi-intensive', which offered the same number of hours distributed in 3-4 months, and the intensive course offered 110 h in 5 weeks during the summer. The results, in contrast to what Bird (2010) found, revealed that concentrating the hours of English instruction in shorter periods of time was more beneficial for the students' learning in certain areas (listening and grammar) than distributing them in many months.

Anderson's ACT models have also greatly relevant discussions on the role of input and output in that they put the emphasis on the relationship between practice type and the transitional change of knowledge. Ellis (1992, 1993) and VanPatten (1996, 2003) take the view that input practice leads to acquisition, but output practice merely serves to improve fluency. On the other hand, studies such as those by DeKeyser (1997), DeKeyser and Sokalski (1996), and Izumi (2002) clearly show a partial lack of transfer between receptive and productive skills (skill-specificity) at the level of proceduralized knowledge but not declarative knowledge.

A number of studies (Benati, 2005; Cadierno, 1995; Erlam, Loewen & Philp 2009; Farley 2004; Van Patten & Sanz, 1995) compared the effect of input based practice with that of output based. Except for the study conducted by Erlam et al., overall other studies found that input-based practice has an impact at least as great as that of the output based.



Erlam et al. (2009) compared the effect of an input based and output based approach to teaching indefinite article for expressing generic meaning. The input group received explicit instruction and structured input practice. The output group received explicit instruction, two types of output practice, namely, text manipulation and text creation, and corrective feedback. It was shown that, using measures of implicit and explicit knowledge, both input and output practice benefited both types of knowledge; however, the output group outperformed the input group. The shortcoming of the study, also mentioned by the authors, was that the output group received the feedback which was different in the nature from what the input group received. Moreover, the output group, unlike the input group, received two types of production practice, which provided two environments with relatively different conditions for proceduralization of the explicit knowledge they received, which might have contributed to their outperformance. Although authors' intention was not to investigate the transferability of skills, the input group's performance on production test was significantly better than that on the pretest. However, they did not administer any measure to see if the output group was also capable of transferring the skill to comprehension context.

Reinders (2009) investigated the effects of three types of production activities on uptake, learners' correct suppliance of the target structure during the treatment, and acquisition, improved performance on a timed and an untimed grammaticality judgment test, of English negative adverbs. The three production activities included a dictation, an individual reconstruction and a collaborative reconstruction activity. It was found that all three activities resulted in uptake, with the collaborative reconstruction and the dictation activities resulting in greater uptake than the individual reconstruction activity. The tasks had an effect on acquisition of grammatical items only. There was no difference between the three tasks on acquisition. It was concluded that a production activity can lead to increased uptake, but not to increased acquisition. For logistic reasons, however, the number of participants in each group was about 15 and the treatment lasted for one week, moreover, there was no control group.

In comparison to other areas related to ACT model and its contribution to SLA, transfer of learning has remained less touched. Even those researchers, who have endeavored to find answers to transfer-of-knowledge- questions, came up with controversial conclusions. Apart from DeKeyser (1997, 2007), who believes in skill-specificity of knowledge, and Van Patten (1996, 2003), who views comprehension activities as sufficient for taking care of both comprehension and production skills, Qin (2008) tried the effect of output and input practice, and concluded that both groups improved in both skills.

Nagata (1998a, 1998b) investigated the effect of computer-assisted production practice and comprehension practice in acquisition of Japanese honorifics and nominal modifiers. The study took



four one-hour sessions over a course of eight days. The post tests were administered two days and the delayed post tests 6 week after the experiment. The results indicated that output computerized practice was more effective than input computerized practice for production and equally for comprehension of both structures. Although the author's focus was not on the knowledge transfer, the results revealed the advantages of output tasks in this regard. However, the results of computer assisted studies cannot be blindly generalized to the contexts where the language teachers have limited access to technological facilities.

Despite large body of data supporting the effect of different practice tasks with certain characteristics on acquiring grammatical structures, research in this area suffers from some methodological shortcomings. First research under laboratory settings is difficult to have implications for ecologically realistic conditions. Most of the studies which addressed the effect of practice type on the acquisition of linguistic elements sufficed with a relatively short period of practice phase. Even those studies within the skill- acquisition strand, have failed to provide practice for a relatively long time to allow the declarative knowledge to become automatized. Except for Seranno et al.'s (2007) study, literature shows that the maximum time allocated for practice was 16 hours (DeKeyser, 2007).

Even fewer studies have attempted to investigate if the skill-specificity is apt to be surmounted by intentional integration or variation of tasks. On the other hand, in spite of having a relatively rich literature on how different tasks might affect language acquisition, there is not much consensus, however, as to the nature, amount, and type of tasks to be used as practice.

Moreover, most of the studies use short retention period, for example one week, 10 days or three weeks at most. Short retention time might make the implementation and generalization of the result troublesome if we consider the conclusion by Driskell, Willis, and Cooper (1992) that "for a cognitive task that has been learnt, retention tends to dissipate after 5 to 6 weeks. This time span seems so long that few educators would consider of any practical value" (p. 771).

Aiming to contribute to filling the gap in empirical research on providing declarative knowledge as conducive (DeKeyser, 2009) to the development of proceduralized grammatical knowledge, the present study intended to examine the effect of form-focused practice tasks when the metalinguistic information on linguistic targets is presented and accompanied by a variety of form focused production tasks to establish proceduralization. The study came up with implications for using tasks in EFL classrooms to improve accuracy and fluency of language learners.



## 2. Research Questions

The present study investigated the following research questions:

1. Does output practice task result in proceduralization of production skill on English modal verbs?
2. Does output practice task result in proceduralization of comprehension skill on English modal verbs?
3. Does output practice task lead into retention of modal procedural knowledge?

## 3. Method

### *3.1. Design*

For the design of the current study a pretest/posttest/delayed posttest intact group design was employed to investigate the effectiveness of output form-focused tasks as practice on the acquisition, retention and transfer of procedural knowledge on English modals.

### *3.2. Context and participants of the study*

The study was conducted at university of Sheikhbahaei in Isfahan, Iran. Two intact classes were involved, making this a sample of convenient. Due to institutional constraints, it was not possible to assign the students randomly to treatment and control groups. At the time of the study all the participants had just enrolled in the second semester. All the participants were, then, freshmen studying English literature, translation or TEFL. They were between 18-25 years of age (average 21), and shared basic demographic characteristics, such as L1, age and field of study.

Since the study addressed the Intermediate population of foreign language learners, the researchers were required to ensure that the participants were homogeneous. To this end, a 60 item Oxford Placement Test, OPT (2001) was administered in the intact classes. Based on the OPT score classification (20-28 elementary, 34-43 intermediate, 50-55- advanced), out of initially 60 participants in two classes 5 were elementary and 2 were advanced. The remaining 52 intermediate participants' score ranged from 34 to 43. Being an intact study, therefore, the elementary and advanced learners also received the treatments. In all the processes of the study the data related to advanced and elementary level learners were excluded from the final results. The total number of participants involved in the study was, then, 52. They were both male ( $N=17$ ) and female ( $N=35$ ). The numbers of participants in the control and output groups were 25 and 27, respectively.



Apart from the intermediate level participants 6 English native speakers participated in the study. They were both male ( $N=4$ ) and female ( $N=2$ ) aged between 32-45, and living in, England, Australia and Canada. The native speakers were all educated at the levels ranging from bachelor degree up to the Ph.D. degrees in different academic majors. Their performance on the tests helped the researchers to modify some items which seemed complicated, ambiguous or difficult for the participants to process. The time limit (see section 3.5) of tests was also decided based on the native speakers' performance.

### *3.3. Target structure*

The target structure in the present study was English modal verbs. A brief review of literature showed that English modals are linguistic features which are cross-linguistically problematic to EFL learners; in terms of developmental properties of L2 acquisition (Pienemann, 1982) modals are acquired early. In terms of when they are introduced in English as a second language courses, modals appear in various levels of pedagogical grading (Ellis, 2009). Although English modal verbs are frequently used in oral and written text, they are non-salient features in that, they tend to be unstressed, hence difficult to notice in the discourse, and their form-meaning mapping is also problematic (Ellis, 2009).

### *3.4. Instructional materials*

The input tasks which served the goal of the present study were:

1. Dictogloss: a structure-based production task, proposed by Wajnryb (1990). It required the use of a short text that was selected or devised to have a structural focuses, modals and subjunctions. The texts were read at normal speed, two times, sentence by sentence, while the learners noted down key words and phrases. The learners then worked personally to reconstruct it as accurately as possible and receive feedback from the researcher particularly on the target structures. (See Appendix A for a sample of this task).
2. Text-reconstruction task: a structure-based production task proposed by Izumi and Bigelow (2000). It required the participants to read a short written passage that was seeded with the target structure and to underline the parts which they felt were important for subsequent reconstruction. The passages were then collected and the participants were asked to reconstruct them as accurately as possible. They were then given the chance to compare what they produced with the original text. (See Appendix A for a sample of this task).
3. Corrected close translation (L1-L2): a form-focused task proposed by Cook (2010). It required learners to produce translation in which they were to keep as close as possible to the original. The



learners then found the opportunity to compare their text with the original one.<sup>3</sup> (See Appendix A for a sample of this task).

### *3.5. Testing materials*

Practice effect on acquisition was operationalized through measuring subjects' performance on two tests: the test which measured proceduralized linguistic knowledge through comprehension skill, Timed Grammaticality Judgment Test, and the test which measured proceduralized linguistic knowledge through production skill, Timed Completion Test. Apart from the timed feature, both of the tests bore structurally irrelevant items to be answered. The features of the tests were decided after Ellis (2009) proposing Time and Focus of attention as two criteria, among others, which are basic to operationalizing the measurement of different types of knowledge, implicit and explicit.

The Timed comprehension/production tests are concerned with learners' being pressured to perform a task online as opposed to when they have an opportunity to plan their response carefully. Operationally, this involves distinguishing tasks that make significant demands on learners' short-term memories and those that lie comfortably within their L2 processing capacity. In other word, timed criterion is to operationalize the degree to which the knowledge has been proceduralized.

Focus of attention is the criterion which is operationalized in two ways: for measuring implicit knowledge the task should require primary focus on meaning, while for measuring explicit knowledge it requires primary focus on form. In the present study Focus of Attention, or lack of it, was operationalized through the inclusion of fillers, irrelevant items in the timed measures of proceduralized knowledge.

The Timed Grammaticality Judgment Test (TGJ) was a pen and paper test consisting of 84 items, grammatical and ungrammatical items. 28 of the sentences were related to the target structure. The remaining 56 items were irrelevant to the target structure so as to minimize the items being form focused. To lower the probability of the test instruction interference with their performance they were given freedom to indicate grammaticality or ungrammaticality of the items in a way they were convenient with. They were asked to underline the wrong part of ungrammatical items. The number of irreverent items was decided to be about two times of the related items, by consulting the related literature.

The time limit for each sentence was established on the basis of native English speakers' average response time, to which was added an additional 20% of the time to allow for the slower processing speed of L2 learners (Ellis, 2009). The mean time spent by the native speakers to answer all the items of TGJT was 17.5 minutes. 3.5 minutes was added to this figure ( $17.5 \times 0.2 = 3.5$ ). The total time



of answering the items was 21 minutes. The average time allocation for each item was, then, about 15 seconds. However, the items were not of equal length, therefore, for short items, they were given about 10 seconds to read and mark them as grammatical/ungrammatical. For longer items they had about 20 seconds to mark the sentences.

The Timed Completion test (TCOM) was a 42-item test. It intended to measure production of the target linguistic feature. There were 14 items related to the target structure. The remaining 28 items were irrelevant so that the participants' attention would be deviated from the target structure. The time limitation and the irrelevant fillers were to make the test a measure of procedural knowledge. The number of irrelevant items was decided to be about two times of the related items, by consulting the related literature.

The time limit for each sentence was established on the basis of native English speakers' average response time for each sentence, to which was added an additional 20% of the time taken for each sentence to allow for the slower processing speed of L2 learners and also producing the missing part (Ellis, 2009). The mean time spent by the native speakers to answer all the items of TCOM test was 12 minutes. 2.4 minutes was added to this figure ( $12 \times 0.2 = 2.4$ ). The total time of answering the items was 14.5 minutes. The time allocation for each item was, then, about 20 seconds. However, the items were not of equal length; therefore, for shorter items they were given about 15 seconds to read and provide the answers. For longer items they had about 25 seconds to provide the response.

In both tests, apart from being asked to read the instruction of the test, on receiving the papers, the participants were also orally explained about what they were expected to do. Although it took quite some time to distribute the papers among the participants, the beginning of the test time was when all the test papers were distributed. On the first page, the test included just some example items and five irrelevant fillers. The inclusion of only irrelevant items on the first page prevented the learners, who received the tests earlier, to start answering the items related to target structures. The participants were given the allocated time for each item by the researcher asking them to move to the next item as soon as the time for answering one item was up. There was a 3 space distance between the items, making control over those who would probably provide the answers in a shorter time and possibly move to the next item. In the completion test the irrelevant items tested various aspects related to verbs of the sentences. This was done since the instruction of the test could not indicate the focuses of the measure. Hence, the participants were required to provide the appropriate form of the verb plus another linguistic item where necessary. For example, some verbs of irrelevant items required the use of preposition; others needed to be presented with appropriate form of *to be* or *to have* and still others depended on providing the appropriate passive structure.



Reliability of the measures, were calculated, using internal consistency of the items in each test over three phases of testing. Cronbach's alpha values for the TGJT in the pre, post and delayed posttest were 0.63, 0.79, and 0.81, respectively. The values for the TCOMT were 0.6, 0.8, and 0.89, in pre, post and delayed posttests, respectively.

### *3.6. Procedures*

A week prior to the administration of the pretests the participants in both groups received a 60 item Oxford Placement Test, OPT (2001). Participants in both groups completed the pretests a week before they received explicit instruction on the target structure. Those students, who answered above 50% of the tests items correct, were excluded from the study. The classes, then, were randomly assigned to one experimental and one control group. Both classes were taught by the same teacher who was also one of the researchers.

In the third week of the study the treatment group received the explicit instruction on the target structure, after that, they received their packet of form-focused tasks which lasted for 12 weeks. The post- tests were administered next week. And the delayed post tests were received five weeks after the administration of the posttests. Table 1 summarizes the procedure of the study.

**Table 1.**

Procedure of the study

Week 1: OPT ( two groups)	
Week 2-Pretests (two groups)	
Week 3- Explicit instruction Output Group)	
Week 4 to15- Output Group	Dictogloss, Text reconstruction, Corrected close translation
Week 4 to 15- Control Croup	Reading comprehension, listening comprehension
Week 16- Posttests ( Two groups)	
Week 21- Delayed posttests ( Two groups)	

As for the explicit instruction the present research adopted the preemptive explicit focus- on- form for instruction of grammar (Ellis, 2007). This type of explicit form-focused instruction is realized by means of metalinguistic explanations. For the present study these typically consisted of information about the property of the target linguistic forms, received by the participants on a handout and explained by the teacher. Further examples followed the oral instruction of the instructor, immediately after the properties were instructed. The explicit instruction took place over one whole session of 90 minutes. The instructor presented the instruction in English and recapped the same



material in Persian to ensure equal understanding. The task sessions were held two times a week. When the explicit instruction was completed the participants received the form-focused tasks in the 4<sup>th</sup> week up to the 15th week. Each session lasted approximately 45 minutes.

The experimental group, overall, had three components: a teacher-fronted explicit explanation of linguistic forms, a combination of three output tasks, and error correction realized as reactive/deductive correction and metalinguistic.

The control were exposed to the same texts, used for the experimental group, yet the questions following texts were comprehension questions and that did not require their focus on the target structure.

A week after the last treatment session, the posttests were administered. In the post and delayed posttests the completion test was administered prior to the grammaticality judgment test.

One version of the tests was used over all three testing phases, pre, post and delayed posttests. However, the order of the items of the tests was different for each test administration.

### *3.7. Scoring and analysis*

For the TGJ each item was allocated a score of either 1(if recognized correctly as grammatical or ungrammatical) or 0 (if incorrectly recognized as grammatical or ungrammatical). The items with no indication of correct/ incorrect, received 0. The mean percentage and standard deviation of each group on pre, post and delayed posttests were calculated.

Responses to the items of completion tests were scored according to three criteria. When they succeeded to provide the appropriate modal together with the correct form of the verb they were scored 1 on the item. When they succeeded in providing the appropriate modal but failed to provide the correct form of the verb they were scored 0.5 on the item. This was done because the response indicated that, at least the participants could recognize which category of modal, probability here, and not ability, should be used. Items with missing and wrong responses were scored 0.

To answer the first two research questions two independent sample t-tests were run to compare the groups' performances on the TGJ and TCOM tests.

The third research question was answered by running two sets of paired-samples t-tests to examine if the effect of form-focused tasks is retained over time. The first set of t-tests compared the



performance of experimental group on pre vs. delayed posttest, and the second set compared that on the post vs. delayed posttests.

#### 4. Results

The groups' mean scores on the TGJ and TCOM pretests were subject to two independent-samples t-tests. The results revealed that there were no significant difference between the two groups at the initial stage of the study: GJ:  $t(50) = -1.3, p = 0.2$ ; CT:  $t(50) = 1.7, p = 0.9$ . This indicated that the groups were equivalent in terms of their ability to comprehend and produce the target structure before the experiment; any difference in the result of post and delayed posttest, then, could be attributed to the effect that instruction and practice might make.

##### *4.1. Production test*

The first research question addressed the proceduralization of production skill as a result of output tasks. The two groups' mean scores and standard deviations on the TCOM pretest, posttest and delayed posttest are presented in Table 2.

**Table 2.**

Descriptive statistics of the groups on the Timed Completion test

Groups	Pretest		Posttest		Delayed posttest	
	M	SD	M	SD	M	SD
Control	21.9	7.8	22.3	8.1	24.6	7.4
Output	16.5	8.6	57.4	16.3	62.3	20.9

As shown in Table 2, the mean percentage score of the output group on the timed completion posttest was better than that of the control group. The figures in the table also indicate that this group's performance on the delayed post test was better than that on the pretest, and even better than the post test.

To see whether the difference between the mean scores of the output group and the control group on the completion posttest was significant an independent samples t-test was run. The result showed that the difference was significant:  $t(50) = -.7.9, p=0$ . The effect size was medium ( $\eta^2 = 0.6$ ), indicating that the magnitude of the difference is meaningful.

The third research question was to investigate whether the probable effect of practice on producing the target structure is retained over time. To this end 2 paired-sample t-tests were run.



The first paired-samples t-test compared the experimental group's performance on the pre vs. delayed posttest of timed completion test. The result of the statistical analysis showed that the difference was significant:  $t(26) = -14, p=0$ . The magnitude of the effect was large (eta squared= 0.9), with the treatment effect explaining up to 90% of the variance.

Another paired-samples t-test was calculated to compare the mean score of the experimental group on the post vs. delayed posttest of timed completion test. The result revealed that the group's performance on the delayed posttest was not significantly different from that on the posttest:  $t(26) = -1.7, p= 0.1$ .

#### *4.2. Comprehension test*

The second research question addressed the proceduralization of comprehension skill. The groups' mean scores and standard deviations on the TGJ pretest, posttest and delayed posttest are presented in Table 3.

**Table 3.**

Descriptive statistics of the groups on the timed GJ test

Groups	Pretest		Posttest		Delayed posttest	
	M	SD	M	SD	M	SD
Control	45.7	10.3	49.2	8.9	43.1	12.7
Output	50	16.1	72.2	11.7	74.3	9.6

As shown in Table 3, the mean percentage score of the output group on the TGJ posttest exceeded that of the control group. The figures also indicate that the output group's performance on the delayed post test was still better than that on the pretest and slightly better than the posttest.

To see whether the difference between the mean percentages of the control group and the output group was significant on posttest, an independent samples t-test was run. The result showed that there was a significant difference between the performances of the groups on the TGJ posttest:  $t(50) = -7.9, p = 0$ . The effect size was relatively large (eta squared= 0.6), indicating that the magnitude of the difference is meaningful.

The third research question was to investigate whether the probable effect of practice on comprehending the target structure is retained over time. To see if the output group has retained its



performance over time two paired samples t-tests were conducted. The first paired-samples t-test compared the group's performance on the pre vs. delayed posttest of TGJ test. The result of the statistical analysis showed that the difference was significant:  $t(26) = -9.5, p=0$ . The magnitude of the effect was large (eta squared= 0.8), with the treatment effect explaining about 80 percent of the variance.

Another paired-samples t-test was calculated to compare the mean score of the group on the post vs. delayed posttest of comprehension. The result revealed that the group's performance on the delayed posttest was not significantly different from that on the posttest:  $t(26) = -1.5, p=0.3$ , indicating that the experimental group has retained its knowledge over time.

## 5. Discussion

The result of pretests of comprehension and production suggested that both groups had initially limited procedural knowledge on English modals. The first and the second research questions addressed the extent to which output practice resulted in acquisition of skill in comprehending and producing the English modals.

The descriptive statistics in Tables 2 and 3 and the results of inferential statistics revealed that output practice of different types enabled the participants to integrate the declarative knowledge into their interlanguage system. The results indicate that on both measures of proceduralized knowledge, where the participants' attention is deviated merely from form through time pressure and irrelevant items, the treatment group were able to process and produce the target structure. As far as the production skill is concerned the experimental group's improvement in their production skill on the posttest was no surprise, since the main purpose of output practice is for the learners to acquire the declarative knowledge and proceduralize it through production tasks. Thus, providing the participants with explicit instruction on the rules and function of English models, followed by a variety of output practice and giving feedback impact significantly on their developing language by helping them attend to the target structure and proceduralize the knowledge. This can be interpreted in light of Swain's Output hypothesis (Swain, 1985). The hypothesis suggests that output practice facilitates acquisition provided that it allows for processes such as noticing, hypothesis testing syntactic processing and metalinguistic reflection. As hypothesized by Swain (1985, production results in the activation of those cognitive processes that lead to restructuring of the developing system). In a similar vein, Levelt's (1989) speech production model might help us interpret the results of the present study. One might hypothesize that output practice gives the learners ample opportunities to conceptualize, formulate and articulate materials in contexts where learners are engaged in the tasks which encourage them to convey message relying on their



existing linguistic knowledge for language production. The cognitive hypothesis of task-based learning also place a heavy responsibility for learning on the necessity of attention and noticing to the L2 structure (Schmidt, 1983, 2001). In accordance with this, the entrance ticket for the knowledge into the developing linguistic system of the learners is attending and noticing the forms: "Learning, establishing new or modified knowledge, memory, skill, and routines is therefore, largely, and perhaps exclusively, a side effect of attended processing" (Schmidt, 2010, p. 4). He further distinguished between "noticing" as limited to the conscious registration of attended instances, and "understanding" as a higher level of awareness that includes generalizations across instances.

As to the acquisition of comprehension skill, i.e., knowledge transfer, it was found that providing learners with a variety of output form-focused practices enables them to acquire the skill to interpret the target structure. In fact introducing a variety of three output-based form-focused tasks as practice in the study has apparently provided opportunities to make form-meaning relationships clear enough to allow the learners to be merely focused, rather than deviated by other requirements of the TGJ test. The result is compatible with the findings in the domain of cognitive and educational psychology where the cognitive psychologists, such as Anderson and Shunn (2000), Schmidt and Bjork (1992), and Speelman and Kirsner (2005) recommended that introducing variation among versions of tasks to be practiced might help the learners generalize their acquired knowledge in the post-training phase to novel situations. Previous studies (DeKeyser, 1997; DeKeyser & Sokalski, 1996), where the skill was practiced through one input and one output task, showed a partial lack of skill transfer to new contexts. The reason that increased variety during the practice stage of learning brings more general skill, according to Speelman et al. (2005) is efficient adaptation to a task environment. When the learners are faced with no or small variation in tasks, they will merely develop skills to cope with that limited environment. When the practice environment involves variety of tasks, even tasks of the same nature, production task in this study, they learn to cope with this variation. What they do under this circumstance is to rely more on the abstraction of features that are common to many task situations. Therefore, they would develop skills to deal with a variety of situations. Anderson et al. (2000) provide an explanation for knowledge transfer which is more related to knowledge-base. They hold that the key to flexible application of knowledge is "understanding", and it is nothing more than having enough knowledge about a concept that can be flexibly used for problem solving involving the concept. This study has, therefore, shown the constructive role of output practice and its variety in transferring the knowledge over context. The TGJ test employed time pressure and irrelevant items to minimize the possibility of the learners monitoring interpretation merely drawing on their declarative knowledge.

This transfer of knowledge, which is called transfer of training in educational psychology, has been addressed by Schmidt and Bjork (1992). They state that introducing some difficulties for the learner



can enhance acquisition as well as retention: "Educators in general and language teachers in particular can induce variation among versions of the tasks to be practiced, with the focus on a criterion of generalizability" (p. 210). Cognitive psychologists have found that for the training of a complex cognitive task, like language learning and use, to be transferred to a new situation, construction of a cognitive structure is inevitable. This enables the learners to recognize particular operations to reach a solution in the new context. Ranzijn (1991) and Shapiro and Schmidt (1982) (cited in Pass & Van Merriënboer, 1994) pointed out that practice variety is beneficial to the construction of cognitive structures and hence to transfer of acquired skill because it increases the chances that similar features can be identified and that relevant features can be distinguished from irrelevant ones.

As to the retention of procedural knowledge for production of the target structure, the result revealed that six weeks after the last session of practice, the participants still performed better than the time they had not received any practice. The practice and the explicit instruction can explain 90% of the variance. Comparing the participants' performance on the TCOM test at the post and delayed post test phase, it was found that the outperformance of the participants did not evade, so that they have not forgotten what they have acquired. Although it is almost impossible to certainly claim whether the knowledge the participants drew upon was implicit/procedural or explicit/declarative, the speed of access to the knowledge benefited the participants after a time span of 6 weeks.

Retention of procedural knowledge for comprehension of the target structure was also shown to persist after a relatively long time. The participants still performed better than the time they had not received any practice. The practice and the explicit instruction can explain 80% of the variance. Comparing the participants' performance on the TCOM test at the post and delayed post test phase, it was found that the outperformance of the output group did not fade away, considering the comparably longer retention time of the delayed post test. This result might be interpreted by resorting to what is called depth of processing. Anderson (2000) cites Craik and Lockhart (1972) who argued that what is critical for the information not to be forgotten is not how long the information is rehearsed but rather the depth to which it is processed. Accordingly, thanks to the communicative nature of tasks the information is rehearsed in the practice phase in a deep and meaningful way. Deep and meaningful processing of information, then, leads to a long-term memory trace (Anderson, 2000).

Another explanation for the maintenance of comprehension ability over time rests on another focal concept in ACT-R theory of skill acquisition. To describe the memory trace, this theory relies on the concept of "activation" which determines the probability of access to memory and the speed with the



memory is accessed. The more a memory is practiced, the stronger it becomes and the more successfully it is retrieved.

Overall, the results of this study suggest that, when institutional limitations do not allow for providing practice of different nature which require different cognitive processes, and definitely requires more time and energy on the part of language teachers to prepare them, introduction of one type of practice might help learners proceduralize the explicit declarative knowledge and transfer that knowledge to relatively dissimilar context of retention test, provided that the teacher introduces variety onto the output tasks.

All in all, the present results confirm that production practices develop production skill as well as comprehension skill. Although the results of this study, and other studies of the same nature, might not guarantee that the knowledge the participants drew on was surely proceduralized, the results confirms what DeKeyser and Juffs (2005) stated about the speeded availability of explicitly acquired knowledge which can be viewed as precious. This knowledge, if not the ultimate goal of language acquisition, can be an intermediate goal on the road to spontaneous language use. Although the real goal of language learning is effortless automatic use of language, automatic use by itself does not exclude the possibility of explicit knowledge accessibility (DeKeyser, 2007). Considering the skill-specificity/generality of knowledge in the skill-acquisition theory, availability of speeded explicit declarative knowledge, or at least partially acquired proceduralized knowledge can enable learners to generalize it to unfamiliar contexts. The pedagogical implication of the study is that besides input practices, there seems to be other options, decided in terms of institutional ecological situations, which can help learners acquire comprehension skills .However, due to some limitations, mentioned bellow, the generalizability of the results requires cautions. Another implication is that at least for structures like modals, for which the mapping of form- meaning relationship is not evident, practices of the same type would lead into a relatively more stable effect.

It should be pointed out that several limitations existed in the process of running the study. To measure the effect of practice on the procedural knowledge, designing instruments which measure the knowledge in an on-line, real communication is felt to be inevitable. However, some structures are less amenable to these types of test; designing a test in which a target structure, like modals, is elicited to occur in obligatory context was, if not impossible, very difficult.



## References

- Anderson, J. R. (1983). *The architecture of cognition*. Cambridge, MA: Harvard University Press.
- Anderson, J. R. (Ed.) (1993). *Rules of the mind*. Hillsdale, NJ: Erlbaum.
- Anderson, J. R. (2000). *Cognitive psychology and its implications*. New York: Worth Publishers.
- Anderson, J. R., & Fincham, J. M. (1994). Acquisition of procedural skills from examples. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 20(6), 1322-40.
- Anderson, J. R., & Fincham ,J. M., & Douglass. S. (1999). The role of examples and rules in the acquisition of cognitive skills. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 23, 932-945.
- Anderson, J. R., & Shunn, C. D. (2000). Implications of the ACT-R learning theory: No magic bullets. In the Glasser (ED), *Advances in instructional psychology*, Vol. 5. (pp. 1-33). Mahwah, NJ: E Erlbaum.
- Benati, A. (2005). The effects of processing instruction, traditional instruction and meaning-output instruction on the acquisition of the English past simple tense. *Language Teaching Research*, 9, 67–93.
- Bialystok, E. (1994). Analysis and control in the development of second language proficiency. *Studies in Second Language Acquisition*, 16, 157- 168.
- Bird, S. (2010). Effects of distributed practice on the acquisition of language English syntax. *Applied Psycholinguistics*. 31, 635-650.
- Byrne, D. (1976). *Teaching oral English*. Harlow: Longman.
- Cadierno, T. (1995). Formal instruction from a processing perspective: An investigation into the Spanish past tense. *Modern Language Journal*, 79, 179-93.
- Carlson, R. A. (2003). Skill learning. In L. Nadel (Ed.), *Encyclopedia of cognitive science*, Vol. 4, (pp. 36- 42). London: Macmillan.
- Cook, G. (2010). *Translation in language teaching*. Oxford: Oxford University press
- Craik, F. I. M., & Lockhart. R. S. (1972). Levels of processing: A framework for memory research . *Journal of Verbal Learning and Verbal Behavior*, 11, 671-684.
- De Graaff , R. (1997). The esperanto experiment: Effects of explicit instruction on second language acquisition. *Studies in Second Language Acquisition*, 19, 249-297.
- DeKeyser, R. M. (1995). Learning second language grammar rules: An experiment with a miniature linguistic system. *Studies in Second Language Acquisition*, 17, 379-410.
- DeKeyser, R. (1997). Beyond explicit rule learning: Automatizing second language morphosyntax. *Studies in Second Language Acquisition*, 19(2), 195–221.



- DeKeyser, R. M. (1998). Beyond focus on form: Cognitive perspectives on learning and practicing second language grammar. In Doughty, C. & J. Williams (Eds.). *Focus on form in classroom second language acquisition* (pp. 42- 63). Cambridge: Cambridge University Press.
- DeKeyser, R. M. (2007). *Practice in a second language: Perspectives from applied linguistics and cognitive psychology*. Cambridge: Cambridge University press.
- DeKeyser, C. (2009). Cognitive-Psychological Processes in Second Language Learning. In M.H. Long and C.J. Doughty (Eds.). *The handbook of language teaching*. (pp. 119-138). Wiley-Blackwell
- DeKeyser, R. M., & Juffs, A. (2005). Cognitive considerations in L2 learning. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (pp. 437-454). Mahwah, NJ. London: Lawrence Erlbaum associates.
- DeKeyser, R. M., Salaberry, R., Robinson, P., & Harrington, M. (2002). What gets processed in processing instruction? A commentary on Bill VanPatten's Processing instruction: An update. *Language Learning*, 52(4), 805-823.
- DeKeyser, R., & Sokalski, K. (1996). The differential role of comprehension and production practice. *Language Learning*, 46(4), 613-42.
- Dornyei, Z. (2009). *The psychology of second language acquisition*. Oxford: Oxford University Press.
- Doughty, C., & Williams, J. (Eds.) (1998). *Focus on form in classroom second language acquisition*. Cambridge, UK: Cambridge University Press.
- Driskell, J. E., Willis, R. P., & Cooper, C. (1992). Effect of overlearning on retention. *Journal of Applied Psychology*, 77, 615-622.
- Ellis, R. (1992). *Second language acquisition and language pedagogy*. Clevedon, UK: Multilingual Matters.
- Ellis, R. (1993). Second language acquisition and the structural syllabus. *TESOL Quarterly*, 27, 91-113.
- Ellis, R. (2002). Does form-focused instruction affect the acquisition of implicit knowledge? *Studies in Second Language Acquisition*, 24(2), 223-236.
- Ellis, R. (2003). *Task-based language learning and teaching*. Oxford: Oxford University Press.
- Ellis, R. (2005). Principles of instructed language learning. *System*, 33, 209-224.
- Ellis, R. (2009). The measurement of implicit and explicit knowledge in Ellis, Loewen, Elder, Erlam, Philp and Reinders (eds.) *Implicit and explicit knowledge in second language learning, testing and teaching*. Bristol, Buffalo, Toronto: Multilingual Matters.
- Erlam, R., Loewen, S., & Philp, R. (2009). Implicit and explicit corrective feedback and the acquisition of L2 grammar. In R.Ellis, S. Loewen, C. Elder, R, Erlam, J. Philp, and H. Reinders (Eds.). *Implicit and explicit knowledge in second language learning, testing and teaching*. (pp. 303-332). Bristol, Buffalo, Toronto: Multilingual matters.



- Ericson, K. A., Krampe, R. T., & Tesch-Romer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *American Psychologist*, 100, 363-406.
- Farley, A. P. (2004). Processing instruction and the Spanish subjunctive: Is explicit information needed? In B. VanPatten (Ed.), *Processing instruction: Theory, research, and commentary* (pp. 227-39). Mahwah, NJ: Lawrence Erlbaum.
- Izumi, S. (2002). Output, input enhancement, and the noticing hypothesis: An experimental study on ESL relativization. *Studies in Second Language Acquisition*, 24(4), 541-77.
- Izumi, S. (2003). Comprehension and production processes in second language learning: In search of the psycholinguistic rationale of the output hypothesis. *Applied Linguistics*, 24(2), 168-96.
- Izumi, S. & Bigelow, M. (2000). Does output promote noticing and second language acquisition? *TESOL Quarterly*, 34, 239-78.
- Levelt, W. (1989). *Speaking: From intention to articulation*. Cambridge, MA: Newbury House.
- Muranoi, H. (2007). Output practice in the L2 classroom. In C. DeKeyser (Ed.). *Practice in a second language: Perspectives from applied linguistics and cognitive psychology*. (pp. 52-84). Cambridge: Cambridge University press.
- Nagata, N. (1998a). Input vs., output practice in educational software for second language acquisition, *Language Learning and Technology*, 1(2), 23-40.
- Nagata, N. (1998b). The relative effectiveness of production and comprehension practice in second language acquisition. *Computer Assisted Language Learning*, 11(2), 153-177.
- Norris, J. M., & Ortega, L. (2000). Effectiveness of L2 instruction: A research synthesis and quantitative meta-analysis. *Language Learning*, 50(3), 417-528.
- Qin, J. (2008). The effect of processing instruction and dictogloss tasks on acquisition of the English passive voice. *Language Teaching Research*, 12(1), 61-82.
- Ranzijn, F.J.A. (1991). The number of video examples and the dispersion of examples as instructional design variables in teaching concepts. *Journal of Experimental Education*, 59, 320-330.
- Reinders, H. (2009). Learner uptake and acquisition in three grammar-oriented production activities. *Language Teaching Research*, 13(2), 201-222.
- Schmidt, R. (1983). Interaction, acculturation and the acquisition of communication competence in M. Wolfson and E. Judd (Eds.). *Sociolinguistics and second language acquisition* (pp. 137-174). Rowley, MA: Newbury House.
- Schmidt, R. (2001). Attention in P. Robinson (Ed.): *Cognition and second language instruction* (pp. 3-32). Cambridge: Cambridge University Press.
- Schmidt, R. (2010). Attention, awareness, and individual differences in language learning. In W.M. Chan, S. Chi, K.N. Cin, J. Istanto, M. Nagami, J. W. Sew, T. Suthiwan, & I. Walker (Eds.), *Proceedings of CLaSIC 2010* (pp. 721-737), Singapore, Centre for language studies.



- Schmidt, R. A., & Bjork, R. A. (1992). New conceptualizations of practice: Common principles in three paradigms suggest new concepts for training. *Psychological Science*, 3(4), 207–17.
- Seranno, R., & Munoz, C. (2007). Same hours, different time distribution: Any difference in EFL? *System*, 35, 305-321.
- Sharwood Smith, M. (1981). Consciousness-raising and the second language learner. *Applied Linguistics*, 2, 159–69.
- Skehan, P. (1998). *A cognitive approach to language learning*. Oxford University Press, Oxford.
- Speelman, C. P., & Kirsner, K. (2005). *Beyond the learning curve: The construction of mind*. Oxford: Oxford University Press.
- VanPatten, B. (1996). *Input processing and grammar instruction: Theory and research*. Norwood, NJ: Ablex.
- VanPatten, B. (2003). *From input to output: A teacher's guide to second language acquisition*. New York: McGraw-Hill.
- Van Patten, B., & Sanz, C. (1995). From input to output: Processing instruction and communicative tasks. In F. R. Eckman, D. Highland, P. Wilee, J. Mileham, & R. R. Weber (Eds.). *Second language acquisition: Theory and language pedagogy* (pp. 169-185). Mahwah, NJ: Lawrence Erlbaum.
- Wajnryb, R. (1990). *Grammar dictation*. Oxford: Oxford University Press.

## Appendix A

Samples of tasks in the output group

### *Dictogloss*

My friend Dona cannot say “no”. If another student asks her, “May I ask for your pen?” she always says, ‘ Of course, you can, and hands it over, even when she only has one pen and it means she cannot do her work. After I heard she did that one day, I told her that she can say ‘Sorry, but you cannot, because I only have one pen’ in reply, she said, ‘But how are they able to do their work without a pen?’ I knew that I could have tried to answer that question, but somewhat I did not think I had the ability to change how she behaved, no matter what I said.

### *Test reconstruction task*

My sister lived in Montana, and when I was visiting her, we would go on weeklong backpacking trips in the mountain. Every morning we woke up to the sound



singing birds. During the day, we used to hike through woods and along mountain streams. Often we saw deer. Once we used to see a bear, but it went off the opposite direction.

*Corrected close translation*

املت درست کردن کار آسانی است. مجبور نیسند آشپز ماهری باشید تا املت درست کنید. دراینجا شما را باچند نکته آشنا مرحله از پخت املت آشنا می کنیم.

۱- اول چند تخم مرغ را در یک کاسه بشکنید. اینکار را با دقت انجام دهید نماید اصلا پوست تخم مرغ داخل املت باشد.

۲- سپس تخم مرغها را مخلوط کنید. مجبور نیستید از غذاساز استفاده کنید.

۳- کمی روغن را در ماهیتابه برشیزید. بهتر است روغن زیتون باشد. اما مجبور نیستید از روغن زیتون استفاده کنید. می توانید از روغن معمولی ذرت هم استفاده کنید. نباید بگذارد که روغن بیش از حد گرم شود.

۴- تخم مرغ های زده شده را در ماهی تابه برشیزید.

۵- وقتی تخم مرغ ها کمی سفت شدند نیمی از آن را روی نیمه دیگر برگردانید. لازم نیست تخم مرغ در این مرحله کاملا سفت باشد.

۶- بعد از چند دقیقه املت را برگردانید. نماید املت بشکند یا له شود.

۷- وقتی املت پخت آنرا بلا فاصله میل کنید. نماید بگذارد خیلی سرد شود چون مزه هی بسیار بدی خواهد گرفت.

## Appendix B

### Timed GJ Test sample of items

1. A- Can't you find that newspaper? B- I am not sure, but, I think someone might have thrown it away. --
2. A- Has the car broken down? B- Well, we may run out of petrol. That's why it doesn't run anymore. --
3. Daughter: I sent my best friend a birthday present, but she never responded or thanked me.  
Mother: I guess, that must get lost in the mail. ---



Timed Completion Test sample of items

1. A: Last night. I heard someone was playing a very difficult piece on the piano. B: I ----- (not be) John. He does not know how to play.
2. A: Oops! I spilled coffee on my shirt. B: You ----- (wash) under hot water before the stain sets.
3. Amy's grammar book is not on her desk. Where is it? B: I don't know. It ----- (be) in her backpack.

### Bio Data

**Mehrnoosh Fakharzadeh** is currently a PhD candidate studying TEFL at the University of Isfahan, Iran. She is an instructor at Sheikhbahaei University, Isfahan, Iran. Her areas of interest are issues at teaching and learning English as a Foreign Language, pragmatics, ESP, and text analysis.

**Manidjeh Youhanaee** holds a PhD in language and linguistics from the University of Essex. She has co-authored with Badrizadeh "A Descriptive Dictionary of Theories of Generative Grammar". She has published a number of articles on the teaching/ acquisition of different English syntactic properties by native speakers of Persian. She teaches MA and PhD courses in linguistics, generative grammar, TEFL and SLA. Her areas of interest include syntactic theory, acquisition of L2 & L3 syntax and issues in teaching and learning English as a second/foreign language.



## Inferential Demands of Different Anaphoric Ties in Expository Texts: Anaphoric Resolution among Undergraduate EFL Learners

Zohreh Yousefvand<sup>1</sup>

*Young Researcher Club, Khorasgan Branch, Islamic Azad University, Khorasgan, Iran*

Ahmad Reza Lotfi

*Islamic Azad University, Khorasgan Branch, Iran*

### Abstract

The readers' ability to integrate current information with given information has been considered as an important component of reading comprehension process. One aspect of this integration process involves anaphoric resolution. The main purpose of the study is to investigate the underlying processes of anaphoric resolution by 90 undergraduate English as foreign language students at Islamic Azad University, Khorasgan Branch (Iran). The study is intended to investigate different types of anaphoric ties occurring in expository texts and their different inferential demands. It has been demonstrated that anaphoric resolution is affected by linguistic and textual constraints and different types of anaphoric ties have different inferential demands. Furthermore, it has been shown that there is a positive relationship between the participants' anaphoric resolution ability and their overall language proficiency level. The findings are also yielded a significant correlation between the ability to resolve one type of anaphoric tie with overall anaphoric resolution ability of the participants.

**Keywords:** Anaphoric devices, Anaphoric resolution, Inferential demands, Expository texts, Language proficiency

### 1. Introduction

Reading skill is crucial since so much of what is needed to be known, especially in academic context, is communicated via the written mode. Accordingly, one of the major concerns of teachers particularly in second/foreign language setting is reading comprehension practice. Teachers know that students need to be good readers in order to comprehend different texts used in their academic courses. In addition, for English as second/foreign language learners (ESL/EFL) in particular, every text that they encounter is a text of English and students who do not read English well will have difficulty understanding different texts in their field of study.

---

<sup>1</sup> Email: z.yousefvand@yahoo.com



Graesser, Millis, and Zwaan (1997) stated that understanding a text consists of constructing a coherent representation of its content. To do so, readers have to understand the relations, which exist between different parts of a text. Furthermore, as stated by Davoudi (2005), a text contains more information than what is stated in a text explicitly. Accordingly, this process of coherent representation of the content of a text also contains representation of information that is expressed by the text implicitly. According to Vonk and Noordman (1990), the computation of this implicit information is referred to as inferencing. This process of inference making is very important in the process of fluent reading and comprehension of the text; hence, if readers are not able to compute inferences that connect explicit information in a text to implicit information, they do not comprehend the text effectively.

An important aspect of this inference making process is anaphoric resolution. This enables the readers to identify anaphoric devices and trace them back by finding the appropriate antecedents, and by doing so integrate new information with given information. According to the Pretorius (2005) definition of anaphor, anaphoric device is a referent relating to a person, entity, event, state, or idea that is introduced into a discourse and again referred to later. Although there is not a unanimous agreement in the literature as to how they should be categorized, the approach taken here in this study is to define five main types of anaphoric devices, namely; repetitions, pronominals, synonyms, paraphrases, and determiners (Halliday & Hasan, 1976). In this study, anaphoric resolution is defined as a reading strategy, which is used by readers to interpret the anaphoric devices that have the same meanings as other elements in the texts (Paterson, Sanford, Moxey, & Dawydiak, 1998; Walsh & Johnson-Laird, 2004). A typical problem in the resolution of anaphoric devices is the presence of more than one candidate for the antecedents of the anaphoric devices. As it is stated by Garrod & Sanford (1994), anaphoric devices vary in terms of their degree of inferential demands, that is, their lexical specificity as well as the extent to which their interpretation is determined by the context. Through reviewing the related literature, the present study supposes that anaphoric devices occur on a continuum based on their hierarchy of difficulty (Kaiser, Runnerb, Sussmanc, & Tanenhausd, 2009).

### *1.1. Language proficiency and anaphoric resolution*

Since the early 1970s, it has been recognized that inferences in general and anaphoric resolutions in particular are crucial aspects of language comprehension. Since then, different aspects of these inferences have been investigated and different studies have tried to investigate the products of inference making. These studies have supported the notion that being able to accurately determine the referents of cohesive devices is important in reading comprehension process (e.g. Kral, 2004). In these studies, the relationship between reading comprehension in L1 and L2 and processing of



specific cohesive ties, such as anaphora has been investigated. The results have revealed that, referential resolution is one of the essential skills to the improvement of students' reading comprehension ability. For instance, Sánchez and García (2009) indicated that correct interpretation of these markers make a significant independent contribution to expository text comprehension under different assessment conditions.

Language proficiency was also shown by a number of studies to affect students' ability to process different anaphoric ties in a written discourse. According to Flanigan (1995), exposure and overall L2 proficiency were the most significant factors in children's ability to comprehend anaphoric references. Pretorius (2005) in her study also revealed that, less skilled readers were less successful in pronominal resolution than skilled readers were. In other word, participants' ability to resolve anaphoric relations was concluded that correlate positively with their overall language proficiency. Megherbi and Ehrlich (2005) also indicated that less skilled comprehenders in a reading situation suffered in the on-line processing of anaphoric pronouns. Their findings showed that less skilled comprehenders were developmentally delayed compared with their skilled peers regarding discourse-level processes.

### *1.2. Problematic anaphoric devices*

In the past three decades, there have been a number of attempts by researchers from different disciplines to discover the nature of anaphoric resolution and its relevance to the process of text comprehension. In the mid-1980s, with progresses in different techniques for measuring online activities different studies have tried to understand about the actual process of anaphoric resolution itself and to understand what readers do as they process the texts.

Particularly, a number of researchers have tried to investigate the problems that readers have in processing different anaphoric devices. They have tried to make it clear that which one of anaphoric ties are most problematic to resolve and the following results have been proposed. Packenham (1980) examined anaphoric resolution among advanced L2 learners of English. He indicated that anaphoric references such as repetitions were easier to resolve than references such as synonyms, which in turn, were easier to resolve than references such as paraphrases. In a related study, Monson (1982) concluded that substitution-ellipsis structures were the most difficult among different anaphoric ties for all age groups. In another study, which was held by Stoefen-Fisher (1988), she found that the null form anaphors were significantly more difficult than the other types of anaphoric relations. In contrast, a number of other studies had shown that in certain discourse contexts, repeated anaphors rather than reduced anaphors might cause difficulty for skilled readers. For instance, Gordon and Chan (1995) and Kennison and Gordon (1997) indicated that anaphoric



resolution was highly sensitive to the focused state of the antecedents, and by focused antecedents it meant the ones that were active in working memory. According to them, reduced anaphoric ties in contrast to repeated anaphoric ties resembled the focused items in a given discourse and as a result were easier to resolve. Al-Jarf (2001) in another study investigated processing of cohesive ties by EFL students. The author in her study attempted to shed some light on the difficulties that EFL college students had in processing different types of inferencing markers. It was found that substitution was the most difficult to process in contrast to repetitions that were the easiest. Pretorius (2005) in her study demonstrated that the anaphoric ties, which involved complex forms of referring expressions such as paraphrases and determiners, were more difficult to resolve.

## 2. Research Questions

The main purpose of the present study was to examine the anaphoric inferencing abilities of EFL readers during reading of expository texts. The study was also intended to investigate the relationship between overall language proficiency level of EFL learners and their anaphoric resolution ability. More specifically the very study was intended to examine the underlying process of anaphoric resolution by undergraduate EFL students. It had also been attempted to investigate the difficulty hierarchy of different anaphoric ties among a special group of EFL students who were native speakers of Persian and to test the utility of hypothesized anaphoric continuum. Accordingly, regarding the present study three broad aims were set up to determine whether;

- There is any relationship between anaphoric resolution ability of undergraduate EFL learners in reading expository texts and their overall language proficiency level.
- There is any relationship between anaphoric resolution ability of EFL learners and different kinds of anaphoric ties occurring in expository texts.
- There is any relationship between the ability to comprehend and resolve one type of anaphoric tie and the students' overall anaphoric resolution ability.

## 3. Method

### *3.1. Participants*

Participants of the study were a group of ninety undergraduate students majoring in English translation at Foreign Language Department of Khorasan Islamic Azad University of Iran. The participants were selected randomly from a pool of accessible population in the mentioned university. The participants were twenty male and seventy female. The age range of the participants was from 20 to 28 with an average age of 24. They were in different academic terms in their



process of studying, namely, 4, 6, and 8. According to their background formal instruction, it was reasonable to assume that the participants would be well familiar with the vocabulary and the structures presented to them.

The participants were categorized into three different groups according to their level of language proficiency. To determine their level of language proficiency an Oxford Placement Test (OPT, 2004 edition) was administered. Based on their OPT scores the participants were divided into three groups. The frequency distribution of the participants in three different groups is shown in Table 1.

**Table 1.**

The frequency distribution of the participants based on their language proficiency level

	Frequency	Range of score	Mean	Std. Deviation
Elementary	29	70-100	88	9.83
Intermediate	32	101-130	120	5.16
Post-Intermediate	29	131-165	141	9.87

The full score of the OPT test was 200. The scores of the participants fell into the range of 70-100 and 101-130 and 131-165, in three different groups. Based on the test classification the participants with language proficiency below 100/200 were identified as elementary foreign language learners, participants whose language proficiency scores were above 130/200 were grouped as post-intermediate foreign language learners, and participants whose scores fell in between were grouped as intermediate foreign language learners.

### *3.2. Procedure*

In order to examine the questions underlying this study a following experiment was carried out. The experiment was originally conducted by Pretorius (2005) who investigated anaphoric resolution among a group of first-year English as second language students. The material for testing anaphoric resolution ability of the participants contained thirty different paragraphs of approximately 72 words [min. 55 / max. 95]. All paragraphs included thirty-eight different anaphoric ties; all of the anaphoric structures were in a backward position (antecedent-anaphor). The frequency distribution of the different anaphoric ties is shown in Table 2.



**Table 2.**

The Frequency distribution of the different anaphoric ties

	Frequency	Percentage
Repetition	8	21%
Pronominal	7	17%
Synonymy	9	24%
Paraphrase	6	16%
Determiner	8	21%
Total	38	100%

As it is tabulated in Table 2 all paragraphs included thirty-eight different anaphoric ties, with an average of eight items per anaphoric category: The anaphoric resolution test included 8 repetitions, 7 pronominals, 9 synonyms, 6 paraphrases and 8 determiners. The paragraphs were taken from different academic textbooks, typical of the texts the students needed to read in an academic context. The anaphoric resolution test had an alpha (Cronbach) reliability score of 0.88.

The first phase of the main study was set to determine the proficiency level of participants through the administration of Oxford Placement Test. The OPT consisted of two sub-tests, namely, listening test and grammar test. Due to the time limitations, the participants took the other task, namely, the anaphoric resolution test. The test on anaphoric resolution included different paragraphs, and all of the paragraphs had underlined anaphoric items. The participants were asked to read each paragraph and then underline all those words/phrases that they thought the underlined words referred back to. Written instructions were provided in English. These instructions were repeated orally in Persian by researcher herself, to avoid any misunderstanding. Although no time limit was set for the test, the time it took for the first five students and for the last five students in the group to complete the test, was noted.

The task of anaphoric resolution was operationalized by asking the participants to underline all the words/phrases they considered as the antecedents of anaphoric ties in different paragraphs. A scoring procedure, which was adopted in the study, was as follows. When the appropriate antecedents of an anaphoric expressions were identified and underlined completely a full mark, (1), was given to each item. When the appropriate antecedents were underlined partially, a half mark, (0.5) was given to each item. Partial underlining referred to cases where, the appropriate antecedents were underlined minimally, the head of the antecedent noun phrases. A zero score, (0), was given to incorrect and/or blank answers.



### *3.3. Analytical tools*

To categorize different anaphoric ties and to determine how the typical EFL readers in the study resolved them, following analytical approach was used. Different types of anaphoric ties were categorized according to traditional Hallidayan text linguistic categories of reference. These different categories were as follows: repetitions, pronominals, synonyms, paraphrases, and determiners.

These five different categories of anaphoric ties were placed on an inference continuum based on their ease of recoverability in terms of their underlying linguistic features. The rationale for this continuum was based on featural overlaps between the antecedents and the anaphors. It was assumed that, the more linguistic features the anaphoric expressions and their antecedents share, the more obvious the link would be, and hence, more easily anaphoric ties would be resolved. On the other hand, the fewer the features were, the less successful the resolution would be (Pretorius, 2005). In the following section, a brief review of these different categories has been presented with illustrating examples.

1. Repetition. The repetition category included anaphoric ties where the anaphor was an exact repetition or a close repetition of an antecedent item. Repetition anaphoric ties were placed first on the continuum, because they contained explicit morphological and semantic clues that could guide the readers to make a link (e.g. believed > belief).
2. Synonym. The synonym category included anaphoric ties where the anaphors were semantically, but not morphologically, related to the antecedents. It was therefore placed second on the anaphoric continuum (e.g. stage > phase).
3. Paraphrase. This anaphoric tie consisted of a determiner this or such followed by a single noun that paraphrased the contents of the antecedent. According to Packenham (1980), the anaphoric item in this category does not have a morphological and/or semantic relationship with an antecedent. Accordingly, paraphrase anaphoric devices were placed third on the continuum (e.g. more challenging work, greater worker participation and control, and more worker autonomy > such conditions).
4. Pronominal. The pronominal tie included pronoun anaphors, which referred to previously mentioned constructs. English pronouns only share gender and number features with their coreferents, and could potentially be linked into several lexical items in a preceding text, as a result they were placed forth on the continuum (e.g. economy food plan > it).



5. Determiner. The determiner category included anaphoric ties where the anaphors comprise only the determiner this (or these). Determiners in English only mark number, therefore provide few morphological and no semantic cues for their antecedents. Consequently, they were placed at the end of the continuum. (e.g. symptoms of patients with brain damage showed, however, that this is not always the case > this). The antecedent of the proceedings' anaphors could referred to a single word or to a more complex concept as expressed in an entire sentence or even a paragraph.

#### 4. Results

The first research question was set to examine the relationship between the L2 proficiency level of participants and their anaphoric inferencing ability. It was hypothesized that students with higher levels of L2 proficiency would be more successful in anaphoric resolution than students with lower levels of L2 proficiency. The participants' mean scores in anaphoric resolution test are shown in Table 3, which demonstrates how anaphoric inferencing abilities are distributed across the different proficient groups.

**Table 3.**

Anaphoric resolution scores across the different proficient groups

Language proficiency level	Mean	N	Std. Deviation	% of Total Sum
Elementary	12.4655	29	4.15316	22.0%
Intermediate	18.8438	32	3.54564	36.8%
Post-intermediate	23.3103	29	4.29371	41.2%
Total	18.2278	90	5.91426	100.0%

The full score of anaphoric resolution test was 38. The mean scores for elementary, intermediate and post-intermediate participants were, 12.46, 18.84 and 23.31, respectively. According to Table 3, the post-intermediate participants outperformed the intermediate ones, which in turn outperformed the elementary participants in anaphoric resolution task. Students performed better in the anaphoric resolution task as the correctness of resolution increased from 22% to about 41%.

The application of a Pearson Product Moment Coefficient Correlation analysis showed a significant correlation coefficient of .740 ( $p < .01$ ) between L2 proficiency and anaphoric resolution. Table 4 presents the results of the Pearson Product Moment Correlation Coefficient between the anaphoric inferencing ability and the language proficiency of participants.



**Table 4.**

The correlation between anaphoric resolution score and language proficiency score of participants

		Language proficiency	Anaphoric resolution
Language proficiency	Pearson Correlation	1	.740**
	Sig. (2-tailed)		.000
	N	90	90
Anaphoric resolution	Pearson Correlation	.740**	1
	Sig. (2-tailed)	.0000	
	N	90	90

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

As shown in Table 4, the anaphoric resolution scores and the language proficiency scores had a positive correlation; when the participants got higher scores in anaphoric resolution task, their scores on language proficiency test were also higher.

A one-way analysis of variance (ANOVA) was further conducted to explore the difference between the three proficient groups regarding their anaphoric inferencing abilities and to examine whether there were significant differences between the three proficient groups. The analysis showed a highly significant effect of language proficiency,  $F (2, 87) = 54$ ,  $p \leq .001$ , indicating that anaphoric resolution were less successful for less proficient participant than for more proficient ones. A post hoc Scheffé test also showed significant differences between different proficient groups with each other (see Table 5).

**Table 5.**

Multiple comparisons using a Post Hoc Scheffé test

(I) language proficiency level	(J) language proficiency level	Mean Difference (I-J)	Std. Error	Sig. 95% Confidence Interval	Lower Bound	Upper Bound
Elementary	intermediate	-6.37823*	1.02439	.000	-8.9295	-3.8270
	Postintermediate	-10.84483*	1.04928	.000	-13.4581	-8.2316
Intermediate	Elementary	6.37823*	1.02439	.000	3.8270	8.9295
	Postintermediate	-4.46659*	1.02439	.000	-7.0178	-1.9154
Post intermediate	Elementary	10.84483*	1.04928	.000	8.2316	13.4581
	Intermediate	4.46659*	1.02439	.000	1.9154	7.0178

\*. The mean difference is significant at the 0.05 level



The results of post hoc Scheff'e test are represented in Table 5. These results provide statistical evidences for differences in anaphoric inferencing abilities in relation to language proficiency level.

The second research question was tried to examine a relationship between anaphoric resolution ability of EFL learners and different kinds of anaphoric ties occurring in expository texts. The first step was to examine the distributions of successful anaphoric resolutions among the five Hallidayan types of anaphoric ties. As it went, it was predicted that ease of resolution would be determined by the amount of morphological and semantic featural overlaps between the antecedents and the anaphoric devices and that successful anaphoric resolution would decrease in this order: repetition > synonymy > paraphrase > pronominal > determiner. Mean overall scores in Table 6 show that undergraduate EFL students in the study were able to identify 44% of the repetitions, 33% of the synonyms, 26% of the paraphrases, 38% of the pronominals, and 52% of the determiners in the texts.

**Table 6.**

Mean overall scores for anaphoric resolution in different categories

	repetition	synonym	paraphrase	pronominal	determiner
Anaphoric resolution	44.044	33.378	26.356	38.556	52.178

Surprisingly, the most successful resolution occurred in the determiner category, which was expected to be the most difficult category. The next category that the students tended to resolve more successfully was the repetition one. Contrary to expectations, the paraphrase category, and not determiner category, proved to be the most challenging one. Consequently, it can be concluded from the results in Table 6 that, paraphrases were the most difficult anaphoric ties under investigation. The second most difficult anaphoric ties were synonyms followed closely by pronominals and repetitions. Determiners were the easiest.

The data were further analyzed in multi-factorial analyses of variance, 5×3 mixed ANOVA, with inferencing (repetition/ pronominal/ synonym/ paraphrase/ determiner) as within-subjects and different proficient group (elementary/ intermediate/ post-intermediate) as between-subjects factors. With regard to between-subjects effects, the results of the previous one-way ANOVA were duplicated. With regard to within-subjects effects, the analysis revealed a significant effect of different categories of anaphoric ties,  $F (4, 87) = 10.28, p <.001$ .

The third question was intended to examine the correlation between comprehending and resolving each one of anaphoric devices and the students' overall anaphoric resolution ability. To examine the



correlations among the five different anaphoric types, a Pearson Product Moment Correlation Coefficient was run. Results of the analysis revealed significant correlation among the repetitions, pronominals, synonyms, paraphrases and determiners mean scores, and between each one of these categories and overall anaphoric resolution ability of participants. The correlations among the five anaphoric categories and between each category and the total score of the other four categories are tabulated in Table 7.

**Table 7.**

Correlation coefficients between the different anaphoric ties categories

	repetition	pronominal	synonym	paraphrase	determiner	total anaph.
repetition	-	-	-	-	-	-
pronominal	.450	-	-	-	-	-
synonym	.471	.434	-	-	-	-
paraphrase	.460	.277	.436	-	-	-
determiner	.416	.410	.569	.350	-	-
total anaph.	.765	.691	.817	.650	.749	-

The correlation matrix shows high and significant correlations between (a) the repetition scores and the total score of the four other anaphoric types ( $r = .765$ ), (b) between the pronominal scores and the total score of the four other anaphoric types ( $r = .691$ ), (c) between the synonym scores and the total score of the four other anaphoric types ( $r = .817$ ), (d) between the paraphrase scores and the total score of the four other anaphoric types ( $r = .650$ ), (e) between the determiner scores and the total score of the four other anaphoric types ( $r = .749$ ). The correlations between the repetitions, pronominals, synonyms, paraphrases and determiners scores were also significant.

## 5. Discussion

Regarding the first question, the results suggested a strong relationship between anaphoric resolution ability of participants and their language proficiency: the post-intermediate students showed a relevant success in anaphoric inferencing test in contrast to students in the intermediate group who in turn inferred anaphors more successfully than those in the elementary group. It had been shown that the more-proficient students were more skillful in identifying anaphoric relations in different paragraphs of the task. Conversely, the less-proficient students connected the anaphoric relations with less success rate. They often referred the anaphors to incorrect or unrelated lexical items or simply skipped them by not answering them. It could be concluded that, the more-proficient students paid more attention to the text elements they read and consequently were able to reach a



more coherent mental representation of information in expository texts. In contrast, the scores obtained by less proficient students for different types of anaphoric relations, suggested that the EFL students who were not proficient language users were not sensitive to the linguistic and textual features of anaphoric ties. Their referring processes were more inconsistent than those of the more proficient readers were and many of their answers were incorrect.

The findings of the study indicated that there is relationship between successful resolution of different types of anaphoric devices and proficiency in English increases. The poor performance of the students in anaphoric resolution could be a result of their poor language proficiency. It can be concluded that, less-proficient EFL readers are not able to perceive morphological and/or semantic similarities between linguistic items (Al-Jarf, 2001). Similarly, it can be argued that the students with low level of language proficiency have low levels of vocabulary and as a result are not able to identify antecedents of anaphoric relations involving synonyms or paraphrases. As Packenham (1980, p. 168) points out, it can be assumed that for successful anaphoric resolution "an adequate knowledge of the meanings of the lexical items used as head nouns in anaphoric NPs or as identifying lexical elements in antecedents" is required. Furthermore, in a good number of studies, vocabulary size has been found to be a good predictor of reading comprehension (e.g. Coady, 1993).

Although the students in the study were not tested for their vocabulary knowledge, it is reasonable to assume that many of them had low vocabulary levels. In this regard, Sharifi (2002) and Talaee (2008) tried to provide an estimate of the vocabulary size of EFL university learners studying at Islamic Azad University Khorasan Branch. The result of their studies indicated that learners possessed a small vocabulary size even in the case of some high frequency words. Furthermore, the result showed an insignificant trend of growth in the vocabulary size of these learners during their years of studies. The results suggested that the trend of growth in the vocabulary size of learners was very slow after semester 4. Of course, further examination of the relationship between EFL students' vocabulary levels and the ease with which they can resolve anaphoric ties is an area that merits more attention.

The second question tried to examine the difficulty hierarchy of different anaphoric ties. According to the findings, a new anaphoric continuum emerged. As a whole, findings indicated that determiners occurred first on the emerged anaphoric continuum, followed by repetitions, pronominals, synonyms and paraphrases. The findings do not support the results obtained in Gordon and Chan (1995), and Kennison and Gordon (1997) studies. Their findings indicated that anaphoric references involving pronominals were easier to resolve than anaphoric references involving repeated nouns. However, in this study, pronominal ties were not resolved most successfully, but they were placed third on the



emerged anaphoric continuum, after determiners and repetitions. Regarding the findings of the present study, it can be argued that resolution of repeated anaphors is sensitive to different factors such as reader's level of proficiency and their stage of reading development. For example, repeated anaphors, which are common in the early grade readers, may be helpful for readers in the early stages of reading and second/foreign language readers whose L2 reading skills is poor (Kennison & Gordon, 1997). On the other hand, except for determiner category, the findings support the results obtained in, Packenham (1980), Monson (1982), Al-Jarf (2001), and Pretorius (2005). They indicated that anaphoric references involving repetitions were easier to resolve and comprehend than references involving synonyms and paraphrases. They concluded that substitution structures such as synonyms and paraphrase were the most difficult structures among different anaphoric ties to be resolved. According to their findings, lexical substitutions (i.e. synonyms & paraphrases) were the most influential in a text's difficulty and readability.

Interestingly, findings of the study indicated that anaphoric ties involving determiners had the highest success rate. The relatively high success rate in resolving determiners suggests that, despite the fact that determiners are potentially opaque and provide few morphosyntactic and no semantic or pragmatic cues for their antecedents, they are very frequent, especially in informative texts and as a result they are familiar anaphoric referents for EFL readers (Fortanet, 2004; Kennison, 2003). The high frequency of determiners in informative texts make them relatively familiar features of written texts, even for less skilled EFL readers, and this may account for their successful resolution. As expected, the paraphrase category of anaphoric relations proved to be a challenging category to resolve. In fact, the most errors, which were occurred in anaphoric resolution test, were those involving paraphrase relations. Paraphrases, particularly those that referred back to longer and complex antecedents, were sometimes ambiguous to resolve, and were difficult even for more-proficient L2 readers of English. This is a factor that writers of expository textbooks should bear in mind, especially when the intended audience comprises EFL readers.

Through examining the correct responses in anaphoric categories, the correlations between resolving different types of anaphoric ties with the students' overall anaphoric resolution ability were investigated. The results of the Pearson Product Moment Correlation Coefficient revealed significant correlations among the repetitions, pronominals, synonyms, paraphrases and determiners mean scores. The findings also revealed significant relationships among mean score of each one of them and total mean score of overall anaphoric resolution ability. Accordingly, it can be hypothesized that, despite of the emerged difficulty hierarchy of different anaphoric types, all of them require somehow the same skills and that mastery of resolving one anaphoric type help in processing the others. Students who are good at identifying and resolving one type of anaphoric type may be good at



others too. To put it in other word, repetitions, pronominals, synonyms, paraphrase and determiners seem to be processed in a somehow similar way.

In conclusion, the present study was intended to shed some light on the difficulties that undergraduate EFL learners had in processing different types of anaphoric ties in reading expository texts. The study also aimed to examine the correlation between ability to resolve different types of anaphoric ties and overall language proficiency level of participants. Anaphoric resolution errors produced by the participants in the study showed that they were unable to make the logical connections between ideas presented in the texts, and thus perhaps were unable to build a coherent mental representation of them. Many of the EFL students who were tested in the study had problems resolving anaphoric devices successfully. Anaphoric resolution became even more unsuccessful when the anaphoric ties involved complex forms of referring expressions such as paraphrases and synonyms.

## References

- Al-Jarf, R. (2001). Processing of cohesive ties by EFL Arab college students. *Foreign Language Annals*, 34, 141-150.
- Coady, J. (1993). Research in EFL/ESL vocabulary acquisition: Putting it in context. In J. Coady, & T. Huckin (Eds.), *Second language learning and vocabulary learning* (pp. 3-23). New Jersey: Ablex Publishing Corporation.
- Davoudi, M. (2005). Inference generation skill and text comprehension. *The Reading Matrix*, 5, 106-123.
- Flanigan, B. (1995). Anaphora and relativation in child second language acquisition. *Studies in Second Language Acquisition*, 17, 331-351.
- Fortanet, I. (2004). The use of 'we' in university lectures: Reference and function. *English for Specific Purposes*, 23, 45-66.
- Garrod, S. C., & Sanford, A. J. (1994). Resolving sentences in a discourse context. In M.A. Gernsbacher (Ed.), *Handbook of psycholinguistic*, (pp. 675-698). New York: Academic Press.
- Gordon, P. C., & Chan, D. (1995). Pronouns, passives, and discourse coherence. *Journal of Memory and Language*, 34, 216-231.
- Graesser, A. C., Millis, K. K., & Zwaan, R. A. (1997). Discourse comprehension. In J. Spence, J. Darley, & D. Foss, (Eds.), *Annual Review of Psychology* (pp. 163-189). Palo Alto, CA: Annual Reviews.
- Halliday, M. A. K., & Hasan, R. (1976). *Cohesion in English*. London: Longman.
- Kaiser, E., Runnerb, J. T., Sussmanc, R. S., & Tanenhausd, M. K. (2009). Structural and semantic constraints on the resolution of pronouns and reflexives. *Cognition*, 112, 55-80.



- Kennison, S. M., & Gordon, P. C. (1997). Comprehending referential expressions during reading: Evidence from eye tracking. *Discourse Processes, 24*, 229-252.
- Kennison, S. M. (2003). Comprehending the pronouns her, him, and his: Implications for theories of referential processing. *Journal of Memory and Language, 49*, 335-352.
- Kral, J. M. (2004). *The effect of pronoun reference drill and practice on the reading comprehension scores of middle school ELLs*. Unpublished master's thesis, University of Saint Paul, Minnesota, United States.
- Megherbi, H. & Ehrlich, M. F. (2005). Language impairment in less skilled comprehenders: The online processing of anaphoric pronouns in a listening situation. *Reading and Writing, 18*, 714-753.
- Monson, D. (1982). *Effect of type and direction on comprehension of anaphoric relationships*. Proceeding of the Annual Meeting of the American Educational Research Association, New York.
- Packenham, K. J. (1980). *The resolution of lexically-reinforced anaphoric reference and its relation to reading in English as a second language*. Unpublished doctoral dissertation, University of Pittsburgh, Russia.
- Paterson, K. B., Sanford, A. J., Moxey, L. M., & Dawydiak, E. (1998). Quantifier polarity and referential focus during reading. *Journal of Memory and Language, 39*, 290-306.
- Pretorius, E. J. (2005). English as a second language learner differences in anaphoric resolution: Reading to learn in the academic context. *Applied Psycholinguistics, 26*, 521-539.
- Sánchez, E. & García, J. R. (2009). The relation of knowledge of textual integration devices to expository text comprehension under different assessment conditions. *Reading and Writing, 22*, 1081-1108.
- Sharifi, M. (2002). Estimating the vocabulary size of Iranian EFL university students. Unpublished master's thesis, Islamic Azad University, Khorasgan Branch, Esfahan, Iran.
- Stoefen-Fisher, J. (1988). Hearing-impaired adolescents' comprehension of anaphoric relationships within conjoined sentences. *Journal of Special Education, 21*, 85-98.
- Talaee, M. (2008). Relationship between vocabulary size and vocabulary learning strategies of EFL learners. Unpublished master's thesis, Islamic Azad University, Khorasgan Branch, Esfahan, Iran.
- Vonk, W., & Noordman, L. G. M. (1990). On the control of inferences in text understanding. In G.B. Balota, D. Flores, & K. Rayner (Eds.), *Comprehension processes in reading* (pp. 94-119). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Walsh, C. R., & Johnson-Laird, P. N. (2004). Co-reference and reasoning. *Journal of Memory & Cognition, 32*, 96-106.



### Bio Data

**Zohreh Yousefvand** is affiliated with Azad University, Isfahan Branch. She is Majoring in the field of TEFL and she has been graduated with honors in her postgraduate studies. She has been an English teacher for more than 8 years. Her scholarly interests range from psychology of language, sociolinguistic to discourse analysis. She has translated different books and also published papers in international journals and conferences.

**Ahmad Reza Lotfi** is assistant professor at Islamic Azad University of Khorasgan, Isfahan.



## Double-track True/False Items: A Viable Method to Test Reading Comprehension?

Purya Baghaei<sup>1</sup>

*English Department, Islamic Azad University, Mashhad Branch, Mashhad, Iran*

Omid Salavati

*English Department, Semnan University, Semnan, Iran*

### Abstract

True-false item format is an objective test form which can be considered as a special type of multiple-choice form. One of the advantages of true-false items is that many can be included in a test and a wide portion of the test content can be covered with them. Furthermore, unlike multiple-choice items true-false items are extremely easy to write. One major disadvantage of true-false items which discourages teachers and test-developers from using them is that they are highly susceptible to guessing. In this paper a scoring procedure named double-track true-false items is suggested to reduce the chances of guessing in conventional true-false items and tested for psychometric appropriateness. Dichotomous and polytomous analyses of a test composed of 40 double-track true-false reading comprehension items show acceptable fit to the Rasch model and a high reliability.

**Keywords:** Double-track true-false items, Multiple-choice items, Rasch model, Objective tests, Test format

### 1. Introduction

Language testing has been through many eras and it has been under influence of different theories of learning and theories of language. One of those influential eras was the psychometric-structuralist at the beginning of the 20th century. The linguistic school of structuralism and the psychological school of Behaviorism had their hands on measurement. In this period, there was a need to identify and measure the learner's mastery of second language proficiency. According to the psychometric-structuralist tenets individual components of language such as phonology, vocabulary, and grammar should have been tested separately (Heaton, 1975). This resulted in a growing interest in discrete-point testing. Reliability and validity were two psychometric concepts which came to fame during this time and were seen as the utmost importance in testing.

---

<sup>1</sup> Email: puryabaghaei@gmail.com



Interest in discrete-point testing and needs of mass testing led to the development of a new type of test method called multiple-choice (MC). Since then, i.e. the heyday of discrete-point testing, multiple-choice tests are still used and they have their own supporters. This popular item format is common among test makers. Everyone has taken multiple-choice tests at one time or another, but probably more often in high school and college than in elementary school (Kubiszyn & Borich, 2003). When facing a multiple-choice question, the student needs to select the correct answer among other alternatives and options which are considered as distractors. Multiple-choice format has some advantages over other formats of testing. Kubiszyn and Borich (2003) list the following advantages to multiple-choice format: (1) Multiple-choice questions have considerable versatility in measuring objectives from the knowledge of the evaluation level, (2) since writing is minimized, a substantial amount of course material can be sampled in a relatively short time, (3) scoring is highly objective, requiring only a count of the number of correct responses, (4) multiple-choice items can be written so that students must discriminate among options that vary in degree of correctness, (5) since there are multiple options, effects of guessing are reduced, and (6) multiple-choice items are amenable to item analysis, which permits a determination of which items are ambiguous or too difficult. Moreover, it is stated that MC format can be easily corrected and it is reliable. However, multiple-choice format has not been without its criticisms and disadvantages. Hughes (2003) mentions these problems for MC format: (1) the technique tests only recognition knowledge, (2) guessing may have a considerable but unknown effect on test scores, (3) the technique severely restricts what can be tested, (4) it is very difficult to write successful items, (5) washback may be harmful, and (6) cheating may be facilitated. Due to the advantages of multiple-choice format and its ease of application in large scale testing, where hundreds of thousands of people are tested, they are used frequently.

Multiple-choice format of testing has been the focus of research for many years. According to Xiaoying Ma (2004), multiple-choice items are more pleasant to test makers because "first, MC items allow for broader content areas to be tested when compared to other formats such as open-ended questions and essays, hence, the use of items generally increases the validity of test score interpretations. Second, sufficient numbers of MC items, if properly constructed, generally produce high reliability of test scores. Last but not least, MC items can be easily pretested, stored, administered, and objectively scored. Thus, the use of MC items can reduce the costs associated with test development, administration, and scoring" (p. 1).

Many studies have investigated different types of multiple-choice format in relation with the number of options and correct responses (Baghaei & Amrahi, 2011; Kubinger & Gottschall, 2007). Item writing is not an easy task. Hence, computer-oriented, sophisticated techniques have been developed to facilitate item writing such as providing the format of the test, adjusting alignment, etc.



They can be also used in correcting the papers. Moreover, there have been some studies providing taxonomies, rules and a few guidelines for constructing multiple-choice items (Haladyna & Downing, 1993; Haladyna, Downing & Rodriguez, 2002). In their taxonomy, Haladyna & Downing (1993) suggest 43 guidelines of which 10 are concerned with general item writing, six are related to stem development, and 20 refer to option development. One of the guidelines concerns the number of options. Haladyna et al. (2002) recommend as many plausible distractors as possible. Reviewing the literature of research on the number of options in MC items, they conclude that the number of options has an impact on item difficulty but not on item discrimination. Generally four or five options would reduce the effect of guessing.

Baghaei and Amrahi (2011a) state, despite “the widespread use of four or five options per item advocated by many authors and test developers, most of the studies carried out to investigate the optimal number of options have ended with recommending the use of three-choice items (p. 193)

A new extension has also been added to multiple-choice format which is called “multiple-response multiple choice” format (MRMC). In contrast with the ordinary MC items which have one correct answer, MRMC items have more than one correct answer requiring test takers to select all the correct answers. According to Hohensinn and Kubinger (2009a) “Conventional multiple-choice formats consist of a single correct answer option and two to seven distractors. Other sorts of multiple choice formats contain more than a single solution which the examinee has to mark. Scoring rules then differ between counting every correctly marked correct answer option and scoring an entire item as solved only if all correct answer options and no distractors have been marked” (p. 231).

According to Xiaoying Ma (2004), “Two types of MR [MRMC] items are commonly used in testing practice, the first has the number of correct answers per item specified to examinees in advance; the other more open-ended type requires examinees to mark all correct answers without mention of the exact number of correct answers.” Since of the main disadvantages of conventional MC format is guessing, MRMC reduces the chances of guessing if a stipulation is made that all correct options and no distractors should be marked to get a point. Some studies have investigated the MRMC format (Kubinger & Gottschall, 2007).

### *1.1. True-false format*

Another objective test format is the true-false (T-F) format. T-F is a type of MC with only two options: true and false. There are a few negative and positive aspects to TF testing format according to Kubiszyn and Borich (2003). The advantages are: (1) because TF questions tend to be short, more



material can be covered than with any other item format. Thus TF items tend to be used when a great deal of content has been covered, (2) TF questions take less time to construct, and (3) scoring is easier with TF questions. To these we can add that T-F items are easily constructed, unlike MC items which are extremely hard to develop.

Kubiszyn and Borich (2003) list the following as the disadvantages of TF format: (1) TF questions tend to emphasize rote memorization of knowledge, (2) TF questions presume that the answer to the question or issue is unequivocally true or false, and (3) TF questions allow for and sometimes encourage a high degree of guessing.

The comparison of T-F format with MC yields some similarities and differences. Both formats are objective and students have to select one option. An MC item has three, four or more alternatives, but a T-F item has only two alternatives. The chances of guessing in an MC item is 25% (assuming there are four options for each item), while in a T-F item it is 50%.

### *1.2. Rational and purpose*

The major discouraging drawback of T-F format is that it is highly guessing-prone. As far as we know, so far no study has investigated a solution to this problem. In this paper, to solve the problem of guessing in TF items, we try to investigate the psychometric qualities of what we call “double-track true-false items.” A double-track TF item has two T-F statements which students have to answer both correctly in order to get the full point for that item. The chances of guessing on a single T-F item are 50%. However, when two T-F statements are put in a single item and stipulated that examinees have to get both right to get a point on the item the chances of guessing is reduced to 25%, as much as in a four-option MC item.

Double-track TF item is not a new method of testing rather it is a scoring rule which rests on the scoring rule of MRMC items where two or more options in an MC item are correct and test-takers are given points if they mark all correct options and none of the distractors (cf. Kubinger).

The purpose of the present study is to test the feasibility of TF items -in terms of psychometric qualities- and the suggested scoring rule to measure reading comprehension of undergraduate English students.



## 2. Method

### 2.1. Participants and Instrument

Eighty true-false reading comprehension items (statements) were written on the basis of four passages. The eighty statements were put in 40 double-track true-false items by putting two independent statements in each item. Three sample items are given below.

1.
  - A. The decayed wooden house was called the 'House of the Rising Sun'.
  - B. Home to the Prince of Darkness is a more fitting description for the house.
2.
  - A. The strange drama started when Mooney's mother reported that he was missing.
  - B. Mooney's dead body was immediately found deep inside a cave.
3.
  - A. The mentally disabled Mooney was brutally beaten to death by mysterious creatures.
  - B. Mooney was allegedly beaten to death by three people-one of them his own mother.

Test-takers were instructed that they should get both statements correct to get a point on an item. The instructions also clearly stated that there was no penalty for guessing. A total of 53 undergraduate students of English as a foreign language at Islamic Azad University, Mashhad Branch took the test. There were 41 females and 12 males among them. The test was given as a final exam test in reading journalistic material at the end of the winter Semester 1390.

### 2.2. Data analysis

The test was scored in two different ways: dichotomous and polytomous. In the dichotomous scoring, a score of 1 was assigned to double-track items in case both statements had been answered correctly. If a test-taker got only one of the statements in an item right the score for the item was 0.

In the polytomous analysis, a score of 2 was assigned when test-takers got both items correct, a score of 1 was assigned when they got only one of the statements right and a score of 0 was assigned when none of the statements in an item had been answered correctly.

The Rasch (Rasch, 1960/1980) dichotomous model was used to analyse the data from the dichotomous scoring and the Rasch particle credit (Masters, 1982) model was used to analyse polytomous scoring data. WINSTEPS (Linacre, 2009) program was used for data analysis.



### 3. Results

#### 3.1. Dichotomous analysis

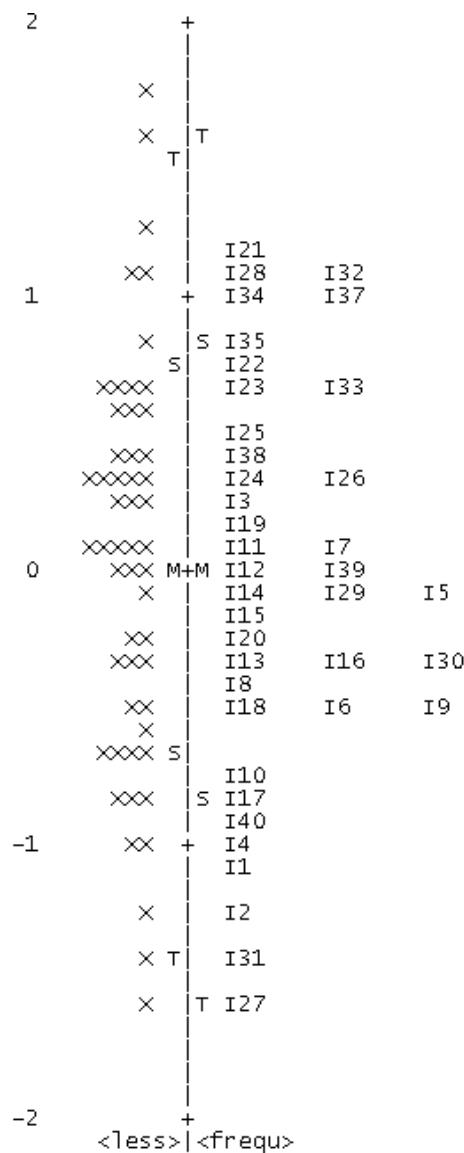
Table 1 shows the item estimates and fit statistics for the 40 items. Items have a difficulty range of 3.86 logit, a reliability and separation of .84 and 2.29, respectively. Persons have a reliability of .75 and separation of 1.71. Person ability estimates range from a low of -1.56 to a high of 1.74, covering an ability span of 3.30 logits, with a standard deviation of .73 and a mean of .03 logits. Figure 1 is a Wright map of items and persons on the variable of interest. The map shows that the items have covered a wide span of the ability continuum, which indicates a precise measurement all over the ability scale.

Fit indices show that three items misfit the model with infit mean square (MNSQ) indices of above 1.30 (Linacre, 2009). Altogether, results show good fit of data to model which is an indication of appropriate internal validity of the test.

**Table 1.**

Item statistics in the dichotomous analysis

ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	ZSTD	OUTFIT MNSQ	ZSTD
1	39	53	-1.10	.33	1.13	.8	1.18	.9
2	40	53	-1.21	.33	1.03	.3	1.07	.4
3	24	53	.24	.29	1.05	.6	1.03	.3
4	38	53	-1.00	.32	.97	-.2	.98	.0
5	28	53	-.10	.29	.94	-.7	.93	-.6
6	33	53	-.53	.30	1.18	1.6	1.18	1.3
7	26	53	.07	.29	.96	-.4	.95	-.5
8	32	53	-.44	.30	.84	-1.7	.89	-.8
9	33	53	-.53	.30	.89	-1.0	.86	-1.1
10	35	53	-.71	.31	.95	-.4	.94	-.3
11	26	53	.07	.29	1.00	-.0	1.02	.2
12	27	53	-.01	.29	.89	-1.3	.90	-1.0
13	31	53	-.35	.29	.97	-.3	.93	-.6
14	28	53	-.10	.29	.87	-1.5	.84	-1.5
15	29	53	-.18	.29	.82	-2.1	.78	-2.1
16	31	53	-.35	.29	.99	0	1.00	0
17	36	53	-.80	.31	.91	-.7	.85	-.9
18	33	53	-.53	.30	.83	-1.6	.78	-1.7
19	25	53	.16	.29	1.07	-.8	1.10	.9
20	30	53	-.27	.29	.85	-1.6	.81	-1.7
21	14	53	1.17	.33	.94	-.3	.95	-.1
22	18	53	.77	.31	1.38	3.0	1.52	2.9
23	19	53	.68	.30	.92	-.7	.94	-.4
24	23	53	.33	.29	.92	-.9	.90	-.9
25	21	53	.50	.30	.93	-.7	.90	-.7
26	23	53	.33	.29	.91	-1.0	.88	-1.1
27	43	53	-1.58	.36	.88	-.5	.72	.9
28	15	53	1.06	.32	1.32	2.1	1.58	2.6
29	28	53	-.10	.29	.91	-1.0	.88	-1.1
30	31	53	-.35	.29	.97	-.2	.93	-.5
31	42	53	-1.45	.35	1.21	1.1	1.65	2.1
32	15	53	1.06	.32	1.32	2.1	1.61	2.7
33	19	53	.68	.30	1.01	1.1	1.06	.5
34	16	53	.96	.31	1.24	1.7	1.33	1.7
35	17	53	.86	.31	.99	0	1.13	.8
36	6	53	2.28	.44	1.14	5	1.55	1.2
37	16	53	.96	.31	1.09	1.7	1.27	1.4
38	22	53	.41	.29	.91	-.9	.87	-1.1
39	27	53	-.01	.29	.86	-1.7	.83	-1.7
40	37	53	-.90	.31	.89	-.8	.85	-.8
MEAN	26.9	53.0	.00	.31	1.00	-.2	1.03	-.1
S.D.	8.5	.0	.80	.03	.14	1.1	.24	1.3



**Figure 1.**

Wright map for the dichotomous analysis



### 3.2. Polytomous analysis

Table 2 shows the item estimates and fit statistics for the 40 items in the polytomous analysis. Items have a difficulty range of 2.76 logits and a reliability and separation of .82 and 2.14, respectively.

Persons have a reliability of .78 and a separation of 1.87. Person ability estimates range from a low of -.06 to a high of 2.84, covering an ability span of 2.90 logits, with a standard deviation of .64 and a mean of 1.27 logits. Figure 2 is a Wright map of items and persons.

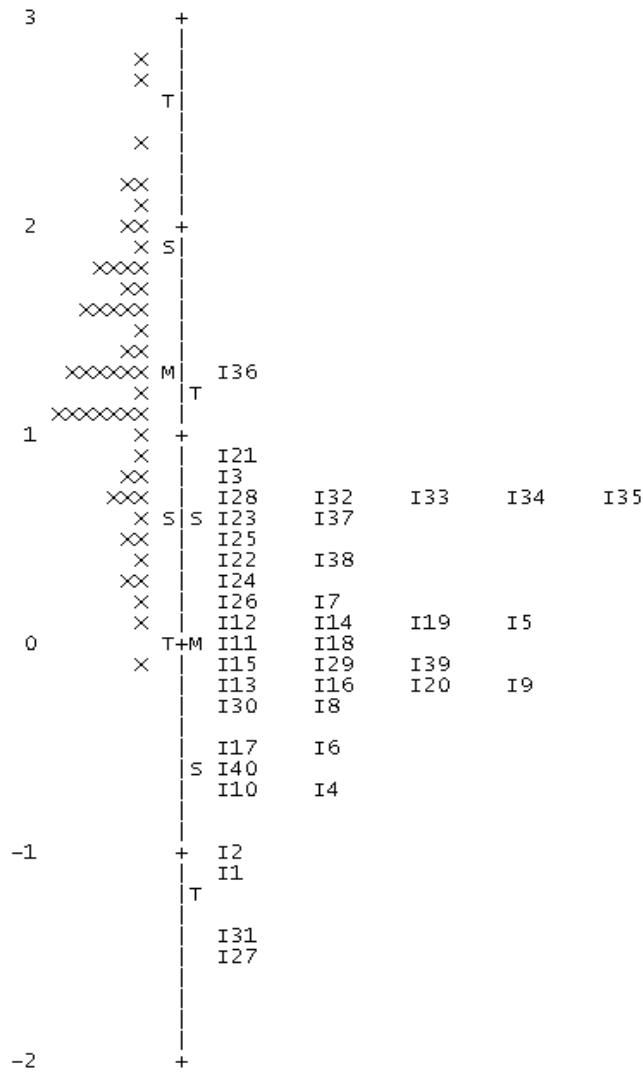
**Table 2.**

Item statistics in the polytomous analysis

ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	ZSTD	OUTFIT MNSQ	ZSTD
1	92	53	-1.07	.31	.99	.0	1.09	.4
2	91	53	-.98	.30	1.31	1.4	1.42	1.6
3	62	53	.82	.23	1.67	3.4	1.63	3.3
4	88	53	-.73	.28	1.15	.8	1.06	.3
5	75	53	.13	.24	1.24	1.4	1.26	1.5
6	85	53	-.50	.27	1.05	.3	1.03	.2
7	74	53	.18	.24	1.00	.1	.96	-.2
8	82	53	-.30	.26	.94	-.3	.95	-.2
9	80	53	-.17	.25	1.23	1.3	1.11	.7
10	88	53	-.73	.28	.83	-.9	.86	-.6
11	78	53	-.05	.25	.80	-1.2	.82	-1.0
12	76	53	.07	.24	.94	-.3	.92	-.4
13	80	53	-.17	.25	1.14	.9	1.07	.4
14	76	53	.07	.24	1.08	.5	1.04	.3
15	79	53	-.11	.25	.87	-.8	.81	-1.1
16	81	53	-.23	.25	1.11	.7	1.10	.6
17	85	53	-.50	.27	1.15	.9	.99	0
18	78	53	-.05	.25	1.40	2.1	1.29	1.6
19	75	53	.13	.24	1.02	.2	1.01	.1
20	81	53	-.23	.25	.86	-.8	.83	-.9
21	61	53	.87	.22	.80	-1.2	.81	-1.1
22	71	53	.35	.23	.94	-.3	1.00	0
23	67	53	.56	.23	.86	-.9	.85	-.9
24	72	53	.30	.23	.85	-.9	.85	-.9
25	69	53	.46	.23	.98	-.1	.96	-.2
26	73	53	.24	.24	.82	-1.1	.83	-1.0
27	96	53	-1.49	.35	.81	-.7	.66	-1.1
28	64	53	.72	.23	1.14	.9	1.15	.9
29	79	53	-.11	.25	.92	-.4	.92	-.4
30	82	53	-.30	.26	1.05	.3	1.02	.2
31	95	53	-1.38	.34	1.10	.5	1.56	1.7
32	65	53	.67	.23	1.07	.5	1.08	.6
33	65	53	.67	.23	1.11	.7	1.10	.7
34	65	53	.67	.23	.95	-.2	.96	-.2
35	64	53	.72	.23	.91	-.5	.89	-.6
36	53	53	1.27	.22	.63	-2.5	.63	-2.5
37	66	53	.62	.23	.80	-1.2	.81	-1.2
38	71	53	.35	.23	.80	-1.2	.80	-1.2
39	79	53	-.11	.25	.79	-1.3	.76	-1.4
40	87	53	-.65	.28	1.18	1.0	1.12	.6
MEAN	76.3	53.0	.00	.25	1.01	.0	1.00	.0
S.D.	9.9	.0	.62	.03	.20	1.1	.21	1.1



Fit indices show that only two items misfit the model. The good fit of data to the polytomous model is an indication of the internal validity of the test when partial credit is given to double-track T-F items half of which are answered correctly.

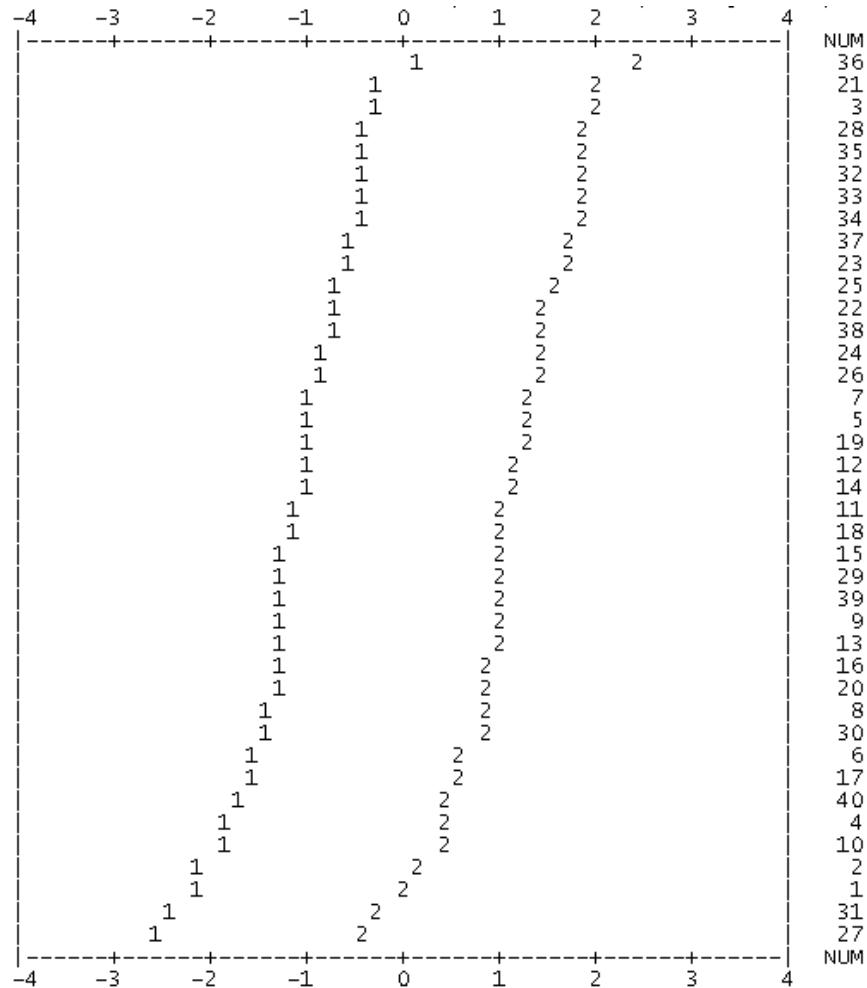


**Figure 2.**

Wright map for the polytomous analysis



Figure 2 is the Wright map for the polytomous analysis. As the map shows the items do not cover the upper part of the ability continuum, where the majority of test-takers fall. However, since in polytomous items there are thresholds associated with each item the operational range of the items is wider than what Figure 2 shows. The range of the items when threshold estimates are taken into consideration is shown in Figure 3. The figure shows that item thresholds range from -2.61 to 2.39 which cover most of the ability span.



**Figure 3.**

Item thresholds map



#### 4. Discussion

The study reported in this article was an attempt to examine the psychometric qualities of true-false reading comprehension items when two T-F statements are put in a single item, namely, double-track true-false items. The procedure can be generalized to any other type of 2-option multiple-choice tests, such as Yes-No items. The purpose of developing this type of items, which is in fact a scoring procedure rather than an item type, was to reduce the effect of guessing on T-F items. The chances of getting a T-F item right is 50% while the chances of getting a 4-option multiple-choice item is 25%. By putting two T-F statements in an item and stipulating that test-takers have to get both statements right to get a point on the item the chances of guessing is reduced to 25%.

Dichotomous Rasch analysis of the 40 double-track T-F reading comprehension items (80 T-F statements), with the stipulation that both statements in an item should be answered correctly to get a point, resulted in three slightly misfitting items with acceptable reliability. The difficulty span of the items was quite wide and covered almost the entire range of the ability scale.

When the items were treated as polytomies, i.e., a score of 1 was given if test-takers had answered one of the statements correctly and a score of 2 was given when they had answered both statements correctly, PCM yielded an equally well-fitting item pool with a slightly higher reliability. Comparing person ability estimates from the two analyses shows that polytomous analysis results in substantially higher ability estimates since credit is given to test-takers when they get half of the items right.

Altogether, the findings show that double-track T-F items, regardless of the method of scoring, have sound psychometric properties as a measure of reading comprehension in English as a foreign language.

Future research should focus on external validation. That is, double-track T-F reading comprehension items should be correlated with more conventional methods of testing reading comprehension such as MC and gap-fill. Examining feasibility of double-track T-F items for the testing of other language skills and components is also worth investigating.

#### References

- Baghaei, P., & Amrahi, N. (2011). The effects of the number of options on the psychometric characteristics of multiple choice items. *Psychological Test and Assessment Modeling*, 53(2), 192-211.



- Haladyna, T. M., & Downing, S. M. (1993). How many options is enough for a multiple choice item? *Educational and Psychological Measurement, 53*, 999-1010.
- Haladyna, T. M., Downing, S. M., & Rodriguez, M. C. (2002). A review of multiple-choice item-writing guidelines for classroom assessment. *Applied Measurement in Education, 15*, 309-334.
- Heaton, J. B. (1988). *Writing English language tests*. Longman: Longman Group UK Limited.
- Hohensinn, C. H., & Kubinger, K. D. (2009). On varying item difficulty by changing the response format for a mathematical competence test. *Austrian Journal of Statistics, 38*(4), 231-239.
- Hughs, A. (2002). *Testing for language teachers*. Cambridge: Cambridge University Press.
- Kubinger, K. D., & Gottschall, C. H. (2007). Item difficulty of multiple choice tests dependent on different item response formats – An experiment in fundamental research on psychological assessment. *Psychology Science Quarterly, 49*(4), 361-374.
- Kubiszyn, T. & Borich, G. (2003). *Educational testing and measurement*. John Wiley & Sons, Inc.
- Linacre, J. M. (2009). WINSTEPS [Computer Software]. Version 3.72.3, Chicago, IL: winsteps.com.
- Linacre, J. M. (2009). A user's guide to WINSTEPS-MINISTEP: Rasch-model computer programs. Chicago, IL: winsteps.com.
- Ma, X. (2004). *An investigation of alternative approaches to scoring multiple response items on a certification exam*. Unpublished Doctoral dissertation. University of Massachusetts. Retrieved from scholarworks.umass.edu/dissertations/AAI3118315/
- Masters, G. N. (1982). A Rasch model for partial credit scoring. *Psychometrika, 47*(2), 149-174.
- Rasch, G. (1960/1980). *Probabilistic models for some intelligence and attainment tests*. Copenhagen: Danish Institute for Educational Research (Expanded edition, Chi-cago: The University of Chicago Press, 1980).

## Bio Data

**Purya Baghaei** is assistant professor in the English Department of Islamic Azad University, Mashhad Branch, Iran. His major research interests are the application of Rasch models in language test development and validation. He was a former data analyst and validation officer at Sprachtestzentrum (Language Testing Centre) at Klagenfurt University, Austria.

**Omid Salavati** is an MA candidate in TEFL program in Semnan University. His studying and research interests are Linguistics, Systemic Functional Linguistics, Neurolinguistics, Cognitive Psychology, Educational Psychology, Critical Discourse Analysis, Conversation Analysis, Second Language Testing & Assessment, Instructed Grammar, and Willingness to Communicate.



## The Effects of Short Message Service- SMS- on University Students' Vocabulary Learning

Minoo Alemi <sup>1</sup>

Zahra Lari

*Sharif university of Technology, Tehran*

### Abstract

Mobile phones as new addition to information and communication technologies have created new ways to help learners in the process of foreign language learning. Given the importance of vocabulary, especially the academic ones for university students, this study tried to investigate the effectiveness of short message service (SMS) on Iranian university students' vocabulary learning and retention. To this end, 28 university students were chosen to take part in the experiment. After administering a pretest and during 16 weeks of experiments, they were taught almost all words of Academic Word List (Coxhead, 2000) via SMS. At the end of the experiment, they were given a vocabulary test, to see the effect of mobile technology on their vocabulary learning and retention. Then, their pre and post test scores were compared using a dependent t-test. The result of t-test for vocabulary test ( $t = 10.63$   $p < .05$ ) showed that their vocabulary knowledge had improved significantly in post test compared to pre test. Also they were administered a delayed post test the result of which showed that besides having a positive effect on vocabulary knowledge, SMS affects students vocabulary retention positively. The result of this study can have pedagogical implication for language teachers, in that they can use SMS as a useful way of teaching large numbers of vocabularies.

**Key words:** MALL, Vocabulary retention, Academic word list, Short message service

### 1. Introduction

Learning new vocabulary is an integral part of learning a new language. Some researchers estimated that in order to understand nontechnical English text, learners should know at least 5000 lexical items (Laufer, 1998; Nation, 2006). Because of the limited class hours, students do not have the opportunity to speak and use all of the words in class. This leads to some problems for language teachers and learners. The problem faced a language teacher is how to teach this large

---

<sup>1</sup> Email: minooalemi2000@yahoo.com



number of vocabulary during the limited time of the class. They can focus and teach individual words explicitly, but deliberately teaching vocabulary is one of the least efficient ways of developing learners' vocabulary knowledge. However it is an important part of a well balanced vocabulary program (Nation, 2005). Also learners face a lot of difficulty in the language learning process. They should memorize and learn the large amount of words of the language they are learning. This urges language teachers and researchers to look for alternative ways to traditional classroom teaching of vocabulary and to change vocabulary learning into a lifelong learning which is not limited to the classroom and adapt it to the needs of learners. They should make learning interesting for students and make them responsible for their own learning. One of the ways that can help teachers in teaching vocabulary is using different technologies available to students. Cognitive and sociocognitive approaches have implication of how to integrate technology in teaching in the communicative trend.

Cognitive psychologists see language learning as a psychological process through which learners construct a mental model of language system based on the interaction of cognitive knowledge and comprehensible input. Some of the technologies that are used in cognitive approaches include text-reconstruction software, concordancing software, telecommunications, and multimedia simulation software (Warschauer & Meskill, 2000).

Sociocognitive approach view language learning as a process of socialization into a particular discourse community. In this perspective, learners should have the opportunity to have interaction in an authentic social context in order to have access to comprehensible input and practice communications in which they will engage in real life (Warschauer & Meskill, 2000). Internet is the technology used in this approach.

One of the technologies that can be used to help learners in learning a foreign language is mobile phones which are dominant in most students' life. They are not just communication devices anymore. They are useful computers that fit into students' pockets, are always with them and nearly always on, and can be used in any kind of learning (Prensky, 2005). This technology has brought about a new type of language learning called Mobile Assisted Language Learning.

The advent of any method of teaching language has always been accompanied by the use of different technologies (Salaberry, 2001) and any act of teaching necessitates using a kind of technology (Warschaer & Meskill, 2000). Mobile assisted language learning is also a by-product of the new approach as the use of technology to teaching that stresses learner centered classroom and learner autonomy. Based on the different definitions provided for mobile learning, it is employing different technologies that makes m-learning different from other kinds of learning



(O'Malley et al., 2003). According to O'Malley et al. (2003) mobile learning includes any kind of learning that take place when the learner is not in a fixed state and take advantage of mobile technology. Generally, according to Geddes (2004) it is the kind of learning which takes place at any time and in any place, that is; it extends teaching and learning outside of the walls of the classroom. Also in this way, students feel responsible for their own learning. Many features of mobile phones – from short message service to recording- have been used to help students to learn a foreign language.

One of the most used features of mobiles in language learning is short message services - SMS. It can provide some of the essential ways to give quick attention to words (Nation, 2001). It is possible to provide a word with its L1 definition, synonym and an example or more in a SMS. Also it is possible to focus on limited amount of information, since too much amount of information can be confusing and discouraging. Besides, they are encouraging to students, because they can study the lessons provided them via SMS any time and in any place that they prefer. Another optimal way of teaching vocabulary which is possible through SMS is spaced presentation which is more effective than massed presentation. According to cognitive psychologists when two presentation of a word are farther apart (spaced presentation), the learners' performance on memory test is significantly better than when the two presentations are close together (Thornton & Houser, 2005).

SMS has been used to help learners in different aspects of language learning especially vocabulary learning. In an attempt to assess the effects of mobile on vocabulary learning, Thornton and Houser (2005) conducted three studies in a Japanese university. First, they surveyed the students regarding their use of mobile phones. The result of the survey revealed that students used their mobiles frequently for sending SMS and emails than talking with their families and friends. Based on this finding, they conducted two counter balanced studies to compare the usefulness of delivering of vocabularies via different mobile media. In one experiment they compared sending some short mini lessons to students' mobiles' emails to presenting them on paper. In another one they compared sending lessons via SMS to presenting them on paper. In the next phase of this experiment they compared the effect of lessons' length on learning. They found no significant difference between long and short lessons. So they concluded that the effect of the frequent study was more important than the quantity of materials presented in a lesson. In another study by Thornton and Houser (2005) a phone sized website was developed to present students some idioms in Japanese and English along with some animation to illustrate them. All in all, these studies showed that mobiles phones can be useful in presenting foreign language materials to students.

In the same line, Lu (2008) used a counter balanced design to investigate the usefulness of short message service (SMS) on 30 vocational high school students' vocabulary retention. In the first



week of a two week experiment, one group received 14 target words via SMS and the other group received the same material on paper. In the second week the two groups changed their media; that is mobile group became paper group and the paper group became mobile group. At the end of each week, an immediate post test, and three weeks later a delayed post test was conducted. The results of the tests showed that both groups regardless of the medium had improved compared to the pretest. But generally mobile group scores were significantly better than the paper group. It's worth mentioning that their scores diminished in the delayed post test. This result underscores the effect of regular reading of vocabularies and it urges to find a way to investigate whether learners read frequently the lessons provided them via SMS. A comparison of students' scores on mobile phone and paper showed that there is a positive relation between reading frequency and vocabulary gain. Also at the end of the experiment the participants were interviewed to find their attitude towards mobile assisted language learning. The information gained by interview showed that generally students had positive attitudes towards mobile vocabulary learning and liked to continue learning vocabulary with the aid of mobile.

In a study conducted in Iran, Derakhshan and kaivanpanah (2011) assessed the effectiveness of SMS on university students' vocabulary learning. During the experiment which lasted for seven weeks, the participants were taught fifteen to twenty words each session and were asked to work in groups to talk about the words. Then the experimental groups were told to send a sentence for each word taught in class to researcher and to three of their classmates. The participants of the control group were asked to write sentences on paper and bring it to the class. The practice of producing output helped to make input comprehensible for students. According to cognitive psychologists having opportunity to produce language helps to make input more comprehensible and learners can construct their own understanding of the language (Warschauer & Meskill, 2000).The result of the post test showed that the experimental group had higher scores than the control group; that is they outperformed the control group in vocabulary retention. Also a delayed post test was administered two weeks later to compare the long term effect of mobile vocabulary learning. It was found that there was no significant difference between the two groups. This result is in line with Lu (2008) in that the participants' scores diminished in the delayed post test and none reached the significant level. However, in both studies, experimental group had more vocabulary gain compared to control group.

Since collocations do not receive enough attention in Iran, Motallebzadeh, beh-Afarin and Daliry Rad (2011) used short message service to help Iranian intermediate EFL learners to retain English collocations. During a five week experiment one group of participants received seventy collocations along with their definitions and some example sentences via SMS and the other group received them on paper. In addition, both groups took two quizzes, one group via SMS and the other on



paper. The result of the post test showed that the participants in the experimental group outperformed those in the control group. Also their scores were significantly better in the post test compared to the pretest. Moreover, according to participants' answers to the attitude questionnaire the participants of the experimental groups were asked to complete, they had positive attitudes towards learning collocations via SMS.

All previously mentioned researches used the short message service feature of the mobile phones to help vocabulary learning. In a different study, Baleghizadeh and Oladrostam (2010) used the recording feature of this technology to assess its effects on learners' grammatical accuracy. During six session of instruction they taught three grammatical categories to participants. Then they chose topics related to those grammatical categories to be discussed in the class and asked the participants of the experimental group to use their mobiles to record their voices while discussing those topics. They had to analyze their speech at home, find their grammatical mistakes and correct them for the next session. At the end of the experiment a grammar test was administered to both groups the result of which showed that the mobile phone helped the experimental group to improve their grammatical accuracy and they performed better than the control group.

The literature shows that most researches have focused on the vocabulary teaching as a whole (e.g. Derakhshan & Kavianpanah 2011; Lu, 2008; Thornton & Houser, 2005) and very few studies employed MALL to teach academic vocabularies. The previously mentioned studies that employed SMS to enhance learners' vocabulary and collocation retention, either used SMS as a supplementary device to classroom vocabulary instruction (Derakhshan & Kavianpanah, 2011), or their experiment lasted for a short period of time and taught students very few number of collocation and words (Mottalebzadeh et al., 2011; Lu, 2008). The present study aimed to assess the effect of SMS vocabulary instruction on university students' academic vocabulary retention.

## **2. Purpose of the Study**

The purpose of this study was conducted to explore the difference between long term and short term retention of words learnt via SMS. In addition, it assessed difference between the short and long term retention of vocabularies learnt via SMS and those learnt by using dictionary. To this end, the following research questions were addressed in this study:

- 1- What is the effect of SMS on university students' vocabulary learning?
- 2- What is the effect of SMS on university students' vocabulary retention?



### 3. Method

#### 3.1. Participants

The participants of this study were 45 non-English major freshman students at Sharif University of technology in Tehran, Iran. One group was assigned as the experimental group ( $N=28$ ) and the other as control group ( $N=17$ ). They were both male and female and were aged between 18-21. They attended general English course two times a week and were taught active skills for reading 4 (Anderson, 2008) by the same instructor. They did not use English outside of class.

#### 3.2. Instruments

One of the instruments which were used in this study was Academic Word List, AWL (Coxhead, 2000). This list is comprised of 10 sublists based on decreasing level of frequency. It totally contains 3000 academic words. The sent words via SMS were chosen from the head words of different sublists. One of the reasons for choosing this list is that it highlights the words that university students may encounter in an academic text, and it accounts for 10% of tokens in the academic cuprous (Coxhead, 2000).

Another reason is the coverage of academic words in participants' text book: *Active Skills for reading 4* (Anderson, 2008). Academic vocabularies are one of the four categories of words identified by Nation (2001). These words have communicative purposes and are essential for students when they take ESP courses. But the research showed that general English books have a low coverage of academic words. According to (Hsu, 2009), AWL accounts for 1.3%-6.54% of the total words in a GE textbook and it accounts for 4.63% of the total words in the participants' text book. While according to Hu (2009) it should have 10% coverage for students to have a successful academic achievement. Hence, the AWL was chosen to be used in this study to compensate for the low coverage of academic words in the textbooks of participants in the experimental and control group in order to help to improve their academic word knowledge.

Another instrument used in this study was a vocabulary test (Appendix). It was composed of 40 multiple choice item which was taken from *Focus on Vocabulary 2: Mastering the Academic Word List* (Schmitt, D. & Schmitt, N., 2001, NY: Pearson Education). It was used as pre test before the beginning of the experiment. Due to the long interval between the pre test and post test, the same test was used as immediate post test at the end of the semester. Also it was used as delayed post test two weeks after the experiment. The reliability of the test was calculated by Cronbach's Alpha which was .89.



### 3.3. Data collection procedures

Before the start of the experiment and in order to assess participants' vocabulary knowledge, they were administered a vocabulary test. During the experiment which lasted for 16 weeks (a complete semester), the participants in the experimental group received 320 vocabularies from the head words of academic wordlist (Coxhead, 2000), their definitions (both Persian and English) as well as example sentences on a regular basis two times a week via SMS. Almost all the words of the AWL (3000 words) were covered in this study. In addition, they were assessed each session by the instructor in class, to see how they learned the vocabularies. The participants of the control group were provided with the same academic vocabularies and were asked to learn them with dictionary by themselves. In addition, like the experimental group they were assessed by the instructor in class. Like the pre and post test, class assessment for both groups was based on the words of the Academic Word List.

In order to see the effects of instruction and assessment, both groups were administered a vocabulary test (immediate post test) at the end of the experiment. Moreover, they took a delayed post test 2 weeks after the experiment to see the long term effect of mobile vocabulary learning on their vocabulary retention.

## 4. Result

To answer the first research question on the effect of SMS on students' vocabulary learning, first the mean scores and standard deviation of the experimental group in the pre and post vocabulary test were calculated and then their pretest and post test scores were compared using a dependent t-test. As shown in Table 1, the participants' mean scores were higher in post test ( $M= 25.17$ ) compared to pretest ( $M= 14.92$ ). This means that the experimental group performed better in post test compared to pretest.

**Table 1.**

Mean and standard deviation of the pre test and post test scores of the experimental group

Group	N	Mean	Std. Deviation
Experimental			
Pretest	28	14.92	4.74
Posttest	28	25.17	5.74



In addition , the result of t-test (Table 2) shows that the SMS had a significant effect on their vocabulary learning; that is, there was a significant difference between their pre and post test scores ( $t= 10.67$ ,  $p= .000$ ).

**Table 2.**

Paired sample t-test result for the pretest and post test scores of experimental group

Experimental group	Paired Differences					t	df	Sig (2-tailed)			
	Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference							
				Lower	Upper						
Pretest posttest	-10.25	5.08	.96	-12.22	-8.27	-10.67	27	.000			

In order to assess the effect of SMS on participants' vocabulary retention (second research question) and compare its short and long term effect, first the mean and standard deviation of the post test and delayed post test scores of the experimental group were computed and then their scores were compared employing a dependent t-test. As seen in Table 3, the experimental group had higher mean scores in delayed post test (26.78) compared to post test (25.17).

**Table 3.**

Mean and standard deviation of the post and delayed post test scores of the experimental group

Experimental group	N	N	Std. deviation
Posttest	28	25.1786	5.74168
Delayed post -test	28	26.7857	5.43212

Table 4 shows the t-test result of the post test and delayed post test of the experimental group. As shown in the table, the experimental group performed significantly better in the delayed post test than the post test ( $t= -2.42$ ,  $p= .02$ ).



**Table 4.**

Paired sample t-test result for the post test and delayed post test scores of the experimental group

Experimental Group	Paired Differences					t	df	Sig. (2-tailed)			
	Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference							
				Lower	Upper						
Posttest – delayed test	-1.60	3.51	.66	-2.96	-.246	-2.42	27	.02			

In order to investigate the difference between the short term retention of the words learnt via SMS and that of the control group who learnt by using dictionary, after computing the mean scores and standard deviation of the post test of the two groups ,their scores were compared using an independent t-test. Table 5 shows the descriptive statistics of the post test scores of the experimental and control groups.

**Table 5.**

Mean and standard deviation of the post test scores of the experimental and control groups

Group	N	Mean	Std. Deviation	Std. Error Mean
Experimental group	28	25.1786	5.74168	1.08508
Control group	23	22.6471	5.18340	1.25716

Table 6 shows the result of the independent t-test of the post test scores of the two groups. The t-test result ( $t=1.48$ ,  $p= .42$ ) and mean scores show that there was not any significant difference between the experimental and control group in the post test. Although there is not a significant difference between the two groups, the effect of text messaging on students vocabulary retention cannot be overlooked. It should be mentioned that before the experiment the control group had more vocabulary knowledge ( $M=16.16$ ) than the experimental group ( $M= 14.92$ ). And after the experiment, it was the experimental group that outperformed the control group, and they had improved more significantly compared to control group.



**Table 6.**

Independent sample t-test result for the post test scores of experimental and control group

	Levene's Test for Equality of Variances		t-test for Equality of Means						
	f	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
								Lower	Upper
Equal variances assumed			1.48	48	.14	2.53	1.70	-.90	5.96
Equal variances not assumed	.65	.42	1.52	38.96	.13	2.53	1.66	-.83	5.89

This result is in line with Derakhshan and Kaivanpanah (2011) who concluded that experimental group had higher mean score than control group in the post test, but the difference between the groups was not significant. But it is not in line with Lu (2008), where not only the experimental group outperformed the control group, but also there was a significant difference between them.

Finally, to explore the difference between the long term retention of vocabularies learnt via SMS and those learnt by using dictionary, first the mean and standard deviation of the delayed post test of the experimental and control groups were computed and then an independent t-test was run to compare the scores of two groups in the delayed post test. Table 7 shows the mean scores and standard deviation of two groups in the delayed post test. As the result shows, both groups have gained higher mean score in delayed posttest compared to posttest and the experimental group ( $M = 26.78$ ) outperformed the control group ( $M= 23.09$ ) in the delayed post test.



**Table 7.**

Mean and standard deviation of the delayed post test scores of the experimental and control groups

Group	N	Mean	Std. Deviation	Std. Error Mean
Experimental group	28	26.78	6.96	1.48
Control group	23	23.09	5.43	1.02

Moreover, the results of the t-test (Table 8) shows that there was a significant difference between the two groups' means scores in the delayed posttest ( $T = -2.109$ ,  $p = .05$ ).

**Table 8.**

Independent t-test result for the delayed post test scores of the experimental and control groups

Levene's Test for Equality of Variances	t-test for Equality of Means								
	f	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
								Lower	Upper
Equal variances assumed									
Equal variances not assumed	1.56	.21	-2.10	48	.040	-3.69	1.75	-7.21	-.17
			-2.04	38.96	.047	-3.69	1.80	-7.34	-.044

Interestingly, this result contradicts the results of previous studies which concluded that both groups' scores decreased in delayed posttest (Derakhshan & Kaivanpanah, 2011; Lu, 2008) and none of the groups reached the significant level.



The findings of this study confirmed the usefulness of SMS on students' vocabulary learning and retention. The results of the present study showed that in the short term, there was not any significant difference between learning vocabulary via SMS and learning it by using dictionary, however, learning vocabulary via SMS helped the participants in the experimental group to enhance their vocabulary knowledge in the post test. But in the long term, vocabulary learning via SMS helped students to retain more vocabulary compared to using dictionary.

This implies that using SMS helps to transfer vocabularies into students long term memory. Since students are used to spending a lot of time using their mobiles, so having words on their mobiles might have encouraged them to review the vocabularies on a more regular basis. This can be proved by the fact that in the delayed post test participants of the experimental group improved significantly and outperformed the participants of the control group.

Despite gaining more score in the post test, it appears that using dictionary did not help the participants in the control group to transfer the learnt vocabularies into their long term memory. Because after two weeks of the experiment and not being assessed by the instructor any more, they did not improve a lot in the delayed post test. It can be due to the fact that learning via dictionary did not lead them to review the vocabularies a lot.

It can be concluded that in order for students to put much effort and time to study new vocabularies they should be responsible for their learning. Moreover, the method of teaching should be so that encourages them to review and study the taught words on a more regular basis and take advantage of students object of interest (in this case mobile phone). Since as it was shown in this study, while both the experimental and control group were responsible for their learning, the experimental (MALL group) group retained more vocabulary in the delayed post test than the control group (dictionary group).

## 5. Discussion

Vocabulary has a central role in language learning, and Language learners should know a large number of words to be successful in their learning or to have a successful communication in it. It also has a vital role in the 4 skills (speaking, reading, writing and listening). Hence this study aimed to investigate the effect of SMS on university students' vocabulary learning and retention. The findings approved the usefulness of SMS in this regard. The participants in the experimental group improved significantly in the immediate post test and outperformed the control group in the delayed post test.



Although being successful in confirming the effect of mobile vocabulary learning on vocabulary retention, this study had some limitations. The first limitation had to do with the number of characters allowed in an SMS. Due to this limitation, only a limited number of words could be sent in each attempt. Also the example sentences were chosen based on this limitation.

Another limitation of the present study was related to the number of participants in the experimental and control group. As was mentioned the experimental group had more participants than the control group. This unequal number might have had an effect on the validity of the results obtained.

Despite these limitations, this study has some implication for incorporation of mobile learning into language courses. The findings can be useful for language teachers and language institutes. It offers them a readily available tool to help improving students' vocabulary knowledge. Since the students are used to using their mobiles and sending and receiving SMS, they can use it as a complementary device to face to face instruction and assessment. In this way they can move toward a leaner centered classroom and make students responsible for their own learning.

This study used mobile learning in a nonreciprocal way, but language teachers can use mobile learning in a reciprocal way. For instance, they can send some tests to students and ask them to send them the answers via SMS.

Like teachers, students can also take advantage of mobile learning. It can help them to learn and retain the large number of the foreign language they are learning .Because the result of the present study confirmed the positive effect on SMS on short and long term retention of vocabulary .

Although this study confirmed the usefulness of mobile learning on vocabulary retention, it did not show its relation to use vocabularies learnt via SMS in extended discourse; more research is needed to explore that whether learning academic words via SMS can help learners to employ those words in extended discourse.

## References

- Baleghizadeh, S., & Oladrostam, E. (2011). The effect of mobile assisted language learning (MALL) on grammatical accuracy of EFL students. *MEXTESOL Journal*, 34(2). 1-10.
- Chen, C. M., & Hsu, S. H. (2008). Personalized intelligent mobile learning system for supporting effective English learning. *Educational Technology & Society*, 11(3), 153-180.



- Chen, C. M., & Li Y. L. (2010). Personalized context-aware ubiquitous learning system for supporting effective English vocabulary learning. *Interactive Learning Environments*, 18(4), 341-364.
- Chinnery, G. M. (2006). Emerging technologies-going to the MALL: Mobile assisted language learning. *Language Learning & Technology*, 10(1), 9-16.
- Clarke, p., Keing, C., Lam, P., & McNaught C. (2008). *Using SMSs to engage students in language learning*. Proceedings of the 20th annual World Conference on Educational Multimedia, Hypermedia & Telecommunications, Vienna, Austria.
- Coxhead, A. (2000). A new academic word list. *TESOL Quarterly*, 34(2), 213-238.
- Deb, S. (2011). Effective distance learning in developing countries using mobile and multimedia technology. *International Journal of Multimedia and Ubiquitous Engineering*, 6(2). 33-40.
- Derakhshan, A., & Kaivanpanah, S. (2011). The impact of text-messaging on EFL freshmen's vocabulary learning. *EUROCALL*, 39-47. Geddes, S.J. (2004) Mobile learning in the 21st century: benefit to learners.  
<http://knowledgegettree.flexiblelearning.net.au/edition06/download/geddes.pdf>.
- Georgieva, E. S., Smrikarov, A. S., & Georgiev, T. S. (2010). Evaluation of mobile learning system. *Procedia Computer Science*, 3, 632-637.
- Hsu, W. (2009). College English Textbooks for general English purposes: a corpus-based analysis of lexical coverage. *Electronic Journal of Foreign Language Teaching*, 6(1), 42-62.
- Hsu, L. (2012). English as a foreign language learners' perception of mobile assisted language learning: a Cross-national study. *Computer Assisted Language Learning*, 1, 17.
- Joseph, S. R. H., & Uther, M. (2009). Mobile devices for language learning: multimedia approaches. *Research and Practice in Technology Enhanced Learning*, 4(1), 7-32
- Kathleen, K. S., & Kitao, K. (2004). TELL and CALL in the millennium: pedagogical approaches in a growing EU-community. *EUROCALL, CONFERENCE REPORTS*.
- Kovalik, C.L., & Hosler, K. A. (2010). Text messaging and the community of inquiry in online courses. *Journal of Online Learning and Teaching*, 6(2), 380-367.
- Kukulska-Hulme, A. & Shield, L. (2008). An Overview of mobile assisted language learning: from content delivery to supported collaboration and interaction. *ReCALL*, 20(3), 271-289.
- Lan, Y. J., Sung, Y. T., & Chang, K. E. (2007). A Mobile-device-supported peer-assisted learning system for collaborative early EFL reading. *Language Learning & Technology*, 11(3), 130-151.
- Laufer, B. (1998). The Development of passive and active vocabulary in a second language: same or different? *Applied Linguistics*, 19(2), 255-271.
- Liaw, S. S., Hatala, M., & Huang, H. M. (2010). Investigating acceptance toward mobile learning to assist individual knowledge management: Based on activity theory approach. *Computers & Education*, 54, 446- 454.



- Lu, M. (2008). Effectiveness of vocabulary learning via mobile phone. *Journal of Computer Assisted Learning*, 24(6), 515-525.
- Motallebzadeh, K. & Ganjali, R. (2011). SMS: Tool for L2 vocabulary retention and reading comprehension ability. *Journal of Language Teaching and Research*, 2(5), 1111- 1115.
- Motallebzadeh, K., Beh-Afarin, R. & Daliry Rad, S. (2011). The Effect of short message service on the retention of collocations among Iranian lower intermediate EFL learners. *Theory and Practice in Language Studies*, 1(11), 1514-1520.
- Mphahlele, M. L., & Mashamaite, K. (2005). The impact of short message service (SMS) language on language proficiency of learners and the SMS dictionaries: a challenge for educators and lexicographers. *IADIS International Conference Mobile Learning*.
- Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge: Cambridge University Press.
- Nation, I. S. P. (2006). How large a vocabulary is needed for reading and listening? *The Canadian Modern Language review*, 63(1), 59-82.
- Nation, P. (2005). Teaching vocabulary. *Asian EFL Journal*, 7(3), 47- 54.
- O'Malley, C., Vavoula, G., Glew, J., Taylor, J., Sharples, M., & Lefrere, P. (2003). *Guidelines for learning/ teaching/ tutoring in a mobile environment*. Mobile project deliverable. <http://www.mobilelearn.org/download/results/guidelines.pdf>
- Prensky, M. (2005). *What can you learn from a cell phone? Almost anything! The innovate gateway* 1 (June/July). Available at: <http://www.innovateonline.info/index.php?view=article&id=83> (last accessed 7 June 2006).
- Salaberry, M. R. (2001). The Use of technology for second language learning and teaching: A retrospective. *Modern Language Journal*, 85, 39-56.
- Sandberg, J., Maris, M., & de Geus, K. (2011). Mobile English learning: an evidence-based Study with fifth graders. *Computers & Education*, 57, 1334-1347.
- Saran, M., Seferoglu, G., & Cagiltay, K. (2009). Mobile assisted language learning: English pronunciation at learners' fingertips. *Eurasian Journal of Educational Research*, 34, 97-114.
- Saran, M., & Seferoglu, G. (2010). Supporting foreign language vocabulary learning through multimedia messages via mobile phones. *H. U. Journal of Education*, 38, 252-266.
- Schmitt, D., & Schmitt, N. (2011). *Focus on vocabulary 2: Mastering the academic word list*. NY: Pearson education.
- Tai, Y., & Ting, Y. L. (2011). Adoption of mobile technology for language learning: teacher attitudes and challenges. *Jaltcalljournal*, 7(1), 3-18.
- Thornton, P., & Houser, C. (2005). Using mobile phones in English education in Japan. *Journal of Computer Assisted Learning*, 21(3), 217-228.
- Trifanova, A., Knapp, J., Ronchetti, M., & Gamper, J. (2004) *Mobile ELDIT: Challenges in transitions from an e-learning to an m-learning system*. Trento, Italy: University of Trento.



Wang, M., & Shen, R. (2011). Message design for mobile learning: learning theories, human cognition and design principles. *British Journal of Educational Technology*, 43(4), 561-575.

Warschauer, M., & Meskill, C. (2000). Technology and second language teaching. In J. W. Rosenthal (Ed.), *Handbook of undergraduate second language education* (pp. 303-318). NJ, Mahwah: Lawrence Erlbaum.

### Bio Data

**Minoo Alemi** is a faculty member of Languages and Linguistics Department at Sharif University of Technology. Her main areas of interest are interlanguage pragmatics, SLA, ESP, and syllabus design. Alemi has published 20 textbooks in general English and ESP, 34 papers in different areas in international journals, and given presentations on TEFL at 57 national and international conferences.

**Zahra Lari** is an M.A. TEFL student in Sharif University of Technology.



## On the Relationship between Personality Types and English Poetry Comprehension

Ali Roohani<sup>1</sup>

Faranak Forouzandeh

*Shahrekord University*

### Abstract

This study aims at investigating Iranian EFL learners' personality profiles, the relationship between personality types and poetry comprehension, the predictors of poetry comprehension and the role of gender in poetry comprehension. To collect data on the personality types and poetry comprehension, Myers-Briggs Type Indicator (MBTI), including 126 dichotomies items, and six English poems, including 180 multiple-choice items, were administered among a sample of 90 Iranian EFL learners. Pearson product-moment correlation, multiple regression, and independent *t*-test statistics were used to analyze the data. Results demonstrated that, first, the judging and perceiving personality types were the most and least reported personality types, respectively. Second, there was no significant relationship between several types of personality and understanding English poetry. Third, there were a significant positive correlation between the judging type of personality and English poetry comprehension and a significant negative one between perceiving type of personality and English poetry comprehension among the participants. Fourth, personality type acted as a predictor of understanding English poetry, but its contribution was not found to be significant. Finally, there was no significant difference between the males and females in terms of poetry comprehension. The pedagogical implications of the findings for second/foreign language (L2) teachers are presented.

**Keywords:** Personality types, Poetry comprehension, Gender, EFL learners

### 1. Introduction

Personality is defined as "the organized, developing system within the individual that represents the collective action of that individual's major psychological subsystems" (Mayer, 2007, p. 14). Personality leads individuals towards dealing with the spotlights of their inner and outer world. In other words, each individual has specific personality to behave in a different way such as when he or she makes decisions. In Myers and McCaulley's (1985) terms, our personality will also lead us to

---

<sup>1</sup> Email: roohani.ali@gmail.com



the place we should focus our attention and get our energy, and to the way we prefer to receive information and make decisions dealing with the outer world. As Myers, McCaulley, Quenk, and Hammer (2003) state, individual differences are present in variations in personalities which are results of some fundamental apparent preferences. As Stake and Eisele (2010) assert, "personality is the study of individual differences," (p. 19). Hence, there is a need to gain a profound understanding of the different ways in which individuals behave. Given the importance of personality, in the last three decades, some studies (e.g., Braden, 1995; Greenberg et al., 2003; Higgins, Peterson, & Rihl, 2007; McPherson, 1999; Noftle & Robins, 2007) have focused on the effect of personality as an inborn trait of everyday life on education. More specifically, these studies have tried to uncover the pivotal role of personality in the learning processes among students in the classrooms. The findings suggest that personality can affect students' behavior in the classroom and academic success, contributing to the relationships students have with their teachers and peers (Matthews, Zeidner, & Roberts, 2006). Given that students should attempt not only to manage their educational and academic goals but also their interpersonal and intrapersonal affairs in life, it is so important for educators to have an awareness of students' personality. As McPherson (1999) states, having the knowledge of student's personality types can assist teachers to match their class schedules and assessments with mixed personalities to achieve high-quality output.

To move further, several recent studies (e.g., Sharp, 2008; Wakamoto, 2007) have studied personality types and their relationship with second/foreign language (L2) acquisition/learning. The results from these studies have shown that L2 learners' personality is an important factor in their L2 learning success and it is associated with L2 learners' strategies on L2 comprehension. In fact, in L2 classrooms, learners are different and have different preferences. It is obvious that some do not enjoy all courses and all types of materials in language classrooms. One rich source of materials for the students and teachers in such L2 classrooms is the literature of the target language students want to learn. Dubash and Anwar (2011) assert that literature as a wordy image of culture which is composed of mores, social behaviors, and knowledge of a group of people should be integrated in L2 learning. In the same line, Eliot (1967, cited in Dubash & Anwar, 2011), a leading English novelist, journalist, and translator, states that "aesthetic teaching is the highest of all teaching because it deals with life in its highest complexity" (p. 9); hence, L2 teaching can be aesthetic if it enjoys literary texts in the L2 classrooms. As Carter and Long (1991) state, literary materials can be a special source for personal development and a means to promote learners' sensitivity, self-awareness, and understanding of the world. As Hayakawa and Hayakawa (1990) point out, poetry as one type of literary text is the superior verbal arts and a kind of authentic materials for L2 learning/teaching. According to Raghupathi (2010), poetry itself can be defined as a kind of language that says more than daily language. Language in poetry is used for its aesthetic qualities in addition to its apparent meaning. Due to the fact that authentic materials are the best for using in



L2 classrooms and help learners achieve L2 in a real world, as Koo (1992) points out, using poetry as a resource provides learners with more opportunities for employing language in an L2 classroom. Providing part of the every language's authentic material and a source of pleasure and enjoyment, poetry persuades us to find some ways to include it in the EFL classes (Moradan, 2006). However, many L2 teachers have not ventured literary materials such as poems as a learning device. In Hess' (2003) terms, poetry as an effective stimulus in L2 classrooms can also bring a kind of participation that other texts cannot make. As Moradan (2006) points out, reading a poem, the reader experiences a particular interaction occurring between him and the poem which is the words and dreams out of the poets mind. The poet, as an individual in the world, may bond to other human kinds by his words in a gorgeous arrangement. During this circle, individuals' personality might determine the poet's intention in his words. Moreover, in an L2 classroom, when an L2 learner come up with problems in comprehending a poem, he or she will participate more effectively in discussions and his or her personality will be involved throughout this participation. It can be thus hypothesized that L2 learners' comprehension of poems may be related with the individual differences such as personality. However, different personalities have different levels of comprehension; teachers can use poems according to their personality types.

In addition, the role of gender in learning cannot be overlooked. Gender as an important dimension in social relations, culture, and individual life can be a significant factor in language learning/acquisition. Several studies (e.g., Connell, 2002; MacKay, & Flege, 2001; Piske, Shehade, 1999) have been done on the effect of gender on learning. Overall, these studies have showed that gender is both an issue of personal identity and a significant social construction (Connell, 2002). According to Padmanugraha (2011), being aware of gender indirectly is helpful for students to live in real world. Since literature is not practical but important in building students' sensitivity to understand others, poetry is the most effective and literary genre to give students this awareness. Consequently, it might be expected that L2 male and female learners may differ in the degree to which they comprehend other texts such as L2 poems too.

In sum, the above issues nudge us towards working on the possible effect of personality type in L2 learning. The related literature is in favor of relationship between L2 learners' personality and L2 learning (Sharp, 2008). However, a need is felt to shed more light on the relationship between personality and comprehending poetry, with an emphasis on the personality theory outlined by Myers and Myers (1995). It is assumed that personality type may relate and contribute to understanding L2 poetry. Also, it is legitimate to explore the role of gender in comprehending L2 poetry since, as Padmanugraha (2011) states, in today's' gendered society poetry may play the role of the most efficient literary genre which can make male and female students aware of building their sensitivity to understand others (Padmanugraha, 2011).



## 2. Theoretical Background

### 2.1. Personality: Myers-Briggs Type Indicator

A typological theory of psychological preferences was first recommended by Jung (1971). As Pittenger (1993) states, becoming interested in type theory after reading Jung's (1921) book, Myers and Myers (1980) attempted to make Jung's theory of human personality more useful in everyday life. These two innovative psychologists began creating an inventory of personality, known as Myers-Briggs Type Indicator (MBTI). They believed if during World War II, women, entering the industrial workforce for the first time knew their personality types; they could participate in the wartime jobs in which they would be successful (Myers & Myers, 1980).

The underlying assumption of personality theory of Myers-Briggs is that "much seemingly random variation in behavior is actually quite orderly and consistent, being due to basic differences in the way individuals prefer to use their perception and judgment" (Myers & McCaulley, 1985, p. 1). According to Myers and Myers (1980), individuals are extravert, introvert, sensing, intuitive, thinking, feeling, judging, and perceiving. Individuals' dissimilar personalities can make them different in the way they focus more, get energy, and prefer to get information and make decisions regarding their outer world.

According to Myers and Myers (1980, 1995), the MBTI was a psychometric questionnaire intended to determine psychological preferences of the individuals' feeling about the world and the way they made decisions in the world. It used four dichotomies combined to create personality types to identify individuals' different preferences. As Callahan (2000) has described, the *Introversion/Extroversion* type represents an individual's energy orientation or primary orientation towards the world; the *Sensing/Intuition* type represents an individual's primary way of gathering information; the *Thinking/Feeling* type describes an individual's primary way of making decisions; the *Judging/Perceiving* type describes the individual's behavior in dealing with the world.

In recent decades, personality has been considered as one of the widely admitted individual differences. Kiany (1998) states that individuals' personalities have lots of effect, generally, on learning and, particularly, on L2 learning/acquisition. For instance, Ehrman and Oxford (1990) studied the language learning styles and strategies of 20 Foreign Service Institute (FSI) college graduate students, the majority of them had studied a west European language, by using the MBTI in a qualitative study. They concluded that introverts, intuitives, feelers, and perceivers seemed to have a learning advantage in the classroom. In another study, Ehrman (1994) chose 831 FSI students. Ehrman discovered that introverts and intuitives achieved higher scores in speaking



proficiency, whereas introverts, intuitives, and thinkers were better readers. Also, sensing types were poor in both reading and speaking.

Later, Carrell (1995) conducted an empirical study about the role personality types played in the writing among writers and raters. She aimed at finding a relationship between writers' personalities and raters' personalities by using the MBTI. Forty-three writers, all native speakers of English, from two sections of a freshman composition class at a major Midwestern university included in the study. Twenty English composition instructors, chiefly doctoral level graduate students, from a major southeastern university made up the raters in the study. They had at least one year's teaching experience in freshman composition. Results showed that the writers' personality types affected the ratings their essays received, and the raters' personality types affected the ratings they gave to the essays. Still, there was no significant relationship between writers' and raters' personality types.

Also, using the MBTI inventory, Callahan (2000) conducted a study in order to explore the relationship between student reflective writing and teacher feedback. She selected three students whose personality types completely differed from her personality as a rater. She analyzed the data and discovered that students as writers needed to go beyond their own preferences and familiarize themselves with other choices. She concluded that the readers who commented on student writings with eagerness could have a critical role in helping students "build on their preferred mental habits and develop their less preferred approaches" (Callahan, 2000, p. 74).

Recently, Wakamoto (2007) tried to find correlations between extraversion and language learning strategies and English listening proficiency. Wakamoto administered the MBTI personality test to 148 female English-major college students in Japan. Strategy Inventory for Language Learning (SILL) was used to measure the strategy choice. Listening proficiency was measured through listening sections of the Test of English for International Communication (TOEIC) and Comprehensive English Language Test (CELT). Extraverts preferred more socio-affective learning strategies. But in teacher-fronted classroom practice, students did not use socio-affective strategies. In addition, he could not prove the effect of extraversion on listening proficiency.

Given the above studies, the MBTI is useful in students writing, and a number of writing teachers suggest administering the instrument in the classroom (DiTiberio & Jensen, 1995; Horning & Sudol, 1997; Thompson, 1995). Also, some researchers (e.g., Callahan, 2000; Carrell, 1995; Thompson, 1995) suggest that the MBTI can be useful in other courses. Consequently, to meet the needs of different types of learners and evaluating their reflections during the course, the MBTI can be used to help in developing teaching methods and reflecting on the various aspects of language learning.



In light of this view, the MBTI is used in the present study to explore L2 participants' personality and find out its relationship with their understanding of L2 poetry.

## *2.2. Poetry*

Emerson (1844, cited in Ahmadzadeh, 2005, p. 2) believes that "language is fossil poetry", the language of the soul; poetry is the black blood out of the poets' mind on behalf of the society. Also, according to Robert Frost, one of the greatest poets, "poetry provides the one permissible way of saying one thing and meaning another" (Rajendra, 1992, p. 7). Whereas literature is a way toward understanding the culture of the target language, many L2 learners are lavished with the technical terms and linguistic expressions, not to mention lack of knowledge about the social context in which many literary pieces are frequently focused on (Finch, 2003). According to Moradan (2006, p. 122), poetry "has a miraculous power," diminishing the tediousness and frustration of our classes. Poetry is a form of literature and can be helpful in L2 learning, but some researchers have barely dealt with it appropriately. Arthur (1968) states we can utilize literature as a potential source in teaching structure, vocabulary, and culture. There are some studies on language learning and poetry even though they are not many in quantity.

For instance, using poetry, Ashton-Hay (2004) in a study of two large classes of Turkish learners in their third and fourth years of the English Language and Literature Department at Ataturk University tried to explore some of the strategies used to challenge, encourage, and support the communicative learning process. The learning strategies covered a range of areas including cognitive reading strategies, rigorous learning standards, learning styles and multiple intelligences, and meaningful content and building cultural knowledge to help learners in recognizing their full potential. Results showed that the suggested learning and teaching strategies had a positive impact on the Turkish undergraduates in Modern English Poetry classes. After conducting these strategies, both classes gradually became more active and communicative in class discussions and enjoyed group work. In conclusion, one can agree with Moradan (2006) who states that poetry "has a miraculous power" (p. 122). Including poetry in the formal education of EFL students can be useful. While reading a poem, the reader faces a particular interaction occurring between him and the poem which is the words the poet's mind. Thus, the poet, as an individual in the world, may bond to other human kinds by his words in a gorgeous arrangement. During this circle, individuals' personality can also determine the poet's intention in his words.

Furthermore, over the last decades, some studies in the social sciences and humanities (e.g., Connell, 2002; Padmanugraha, 2011) have concentrated on the notions of masculinity and femininity. Some of these studies, in general, consider the role of gender as a dynamic motivating



factor and suggest that gender is a significant dimension of social relations, culture, and individual life. Until 1990, there have been quite a few empirical studies on gender and L2 learning, but there has been growing interest in gender differences recently. For instance, Shehade (1999) found out that men exploited the conversations to enhance their performance or production skills in L2 conversation classrooms whereas women used the conversations to enhance their comprehension skill in L2 conversation classrooms. Piske, MacKay, and Flege (2001) studied the effect of Italian-English bilinguals' gender on the degree of L2 accent. In contrast, they reported that gender had no significant and independent effect on the L2 pronunciation accuracy. The results are not consistent.

As Stake and Eisele (2010) has mentioned, "personality is the study of individual differences," (p. 19). Hence, it helps individuals to understand the way our gendered society forms and underlines the differences in women's and men's feelings and behaviors better; bearing gender in mind and its association with personality has some implications for L2 teaching and learning. Besides, the foregoing review of the literature shows that there is a dearth of research on the relationship between L2 learners' personality types and understanding L2 poetry, the role of gender in comprehending L2 poetry and unearthing the predictors of L2 poetry comprehension. The present study seeks to shed more light on the issues at hand to help L2 instructors to develop teaching methods and techniques so that they can meet the needs of learners with different personality types. This study, thus, seeks answers to the following questions:

1. What are the Iranian EFL learners' personality profiles?
2. To what extent is there any relationship between personality type and understanding English poetry among Iranian EFL learners?
3. What can act as the best predictor of poetry scores: personality type or gender?
4. Is there any significant difference in English poetry comprehension between male and female Iranian EFL learners.

### **3. Method**

#### *3.1. Participants*

A total of 90 students (30 males and 60 females) majoring in Teaching English, English literature and translation at Shahrekord university and University of Isfahan were participated in the study. The ratio of females to males was two to one.



### 3.2. Instrumentation

The following instruments were used in this study to collect data:

#### 3.2.1. The Myers-Briggs Type Indicator (MBTI)

The Myers-Briggs Type Indicator (MBTI) was used in the current study to identify personality type and determine the personality profile of the participants. The MBTI has been published in three forms: (1) Form F, the research form with 166 items, (2) Form G, the standard form with 126 items, and (3) Form H or AV, the abbreviated version, a self-scoring form with 50 items. Form G of the MBTI, which is a 126-item pencil-and-paper inventory, was used in current study (Myers & McCaulley, 1985). In addition, according to Myers and McCaulley (1985), all forms of the MBTI have been widely tested for validity and reliability, so they are quite accurate measurements appropriate for the adult populations; the reliability coefficients for the four preference indexes of the MBTI ranged from .95 for the Extraversion-Introversion (E-I) Index to .96 for the Sensing-Intuition (S-N) and Judgment-Perception (J-P) Indexes, to .98 for the Thinking-Feeling (T-F) Index. Besides, its reliability and validity weighed favorably against other personality instruments too (Sundberg, 1970). In this study, the reliability of the instrument was found to be 0.75 through Cronbach's alpha, demonstrating that the instrument was reliable to be used in this study.

#### 3.2.2. Poetry test

To collect data on the participants' understanding of the poetry, six English poems were selected and 180 multiple-choice items were developed. The poems selected included: *The River*, by Yi Yung (n.d.), *City Life* by D. H. Lawrence (1885-1930), *A Travel Poem* by Teika (13<sup>th</sup>), *With How Sad, O Moon Thou Climb'st* by Sir Philip Sidney (1554-1586), *A Winter Poem* by Teika (13<sup>th</sup>), and *Under the E-Train* (anonymous). Following Bachman (1990), the validity of the poetry test was accounted for through checking both test development and test use, particularly investigating the table of test specifications, by two Iranian assistant professors who were non-native speakers of English, but had adequate knowledge of English poetry and were considered as experts in the area of English poetry.

### 3.3. Procedures

Before carrying out the main study, ten students resembling those in the main study participated in a pilot study to find out the appropriacy of the items. Since the number of the participants was small,



the test-retest estimate for the poetry test was obtained. The reliability estimate in the pilot study was .075, so it was acceptable with a small sample.

The selected poems were administered in one session to the participants from the two aforementioned universities during the 2010-2011 Academic Year. After the poetry test had been administered, the MBTI was administered to the same participants in a separate session to determine the personality profiles of the participants.

Finally, SPSS (version 17.00) was used to conduct descriptive statistics, correlational analysis, multiple regression, and *t*-test so as to address the research questions of the study.

#### 4. Results

To answer the first research question, descriptive statistics were obtained for all types of personalities. Table 1 represents the descriptive statistics of the eight types of personalities.

**Table 1.**

Descriptive statistics of eight types of personalities

Variable	N	Min	Max	Mean	Std. Dev.
Extroverted	90	22	79	50.53	12.44
Introverted	90	21	78	49.47	12.44
iNtuitive	90	22	72	45.65	9.85
Sensing	90	28	78	54.35	9.85
Thinking	90	26	76	54.58	10.44
Feeling	90	24	74	45.42	10.44
Judging	90	31	78	58.24	12.04
Perceiving	90	22	69	41.76	12.04

As demonstrated in Table 1, the mean scores ranged from 41.76 to 58.24, indicating that the performance of the participants on the dichotomy of judging/perceiving varied much. This means that, most of the Iranian EFL participants possessed the judging personality ( $M = 58.24$ ). Furthermore, the least common type was the perceiving personality ( $M = 41.76$ ). As depicted in Table 1, the participants perceived themselves stronger at judging type ( $M = 58.24$ , Std. Dev. = 12.04), followed by thinking type ( $M = 54.58$ , Std. Dev. = 10.44) and sensing type ( $M = 54.35$ , Std. Dev. = 9.85).



Judging/perceiving dichotomy shows the way individuals manage the outside world. According to Oxford (2003), judging learners having a tendency to a planned and structured life enjoy printed instruction and particular assignments with deadlines and are meticulous in learning process. The high mean score of the judging personality ( $M = 58.24$ ) in the sample suggests that the participants of the current study had a tendency to be more judging than the other type. In contrast, considering other individuals' feelings, perceiving learners like to discuss and learn a lot through this kind of negotiations (Razawi, Muslim, Che Razali, Husin, & Abdul Samad, 2011). However, in the Iranian EFL context of the present study, most of the L2 learning classroom settings are traditional. Thus, as Oxford (2003) argues, perceiving learners are in unfavorable circumstances in such contexts and it is difficult for them to sprout and grow. Thereby, the abovementioned results indicate the L2 participants were more judging and likely to learn an L2 in a planned and systematic order. On account of the fact that in Iran L2 teaching is typically grammar-based (Rahimi, 2005), Iranian L2 learners often tend to learn English seriously and in a structured way.

To address the second research question the Pearson product-moment correlation coefficients were obtained between eight types of personalities and comprehending English poetry scores. Besides, preliminary statistics were performed to ensure no violation of the assumptions of normality, linearity, and homoscedasticity.

**Table 2.**

Correlation between eight types of personality and poetry comprehension

Variables	E	I	N	S	T	F	J	P
Poetry	-.206	.206	.019	-.019	.111	-.111	*.354	*-.354
Comprehension	.150	.150	.895	.895	.442	.442	.012	.012

*Note.* \*. Correlation is significant at the 0.05 level (2-tailed); E = Extroverted; I = Introverted; N = iNtuitive; S = Sensing; T = Thinking; F = Feeling; J = Judging; P = Perceiving.

As depicted in Table 2, there was a significant and medium positive correlation between poetry comprehension and judging personality,  $r(90) = .35$ ,  $*p < .05$ , with high levels of judging personality associated with high scores in poetry comprehension. Moreover, there was a significant and medium negative correlation,  $r(90) = -.35$ ,  $*p < .05$ , between the perceiving personality and L2 participants' poetry comprehension. Also, the correlation between the introvert, intuitive, and thinking personality types and poetry compression was positive, but not significant. In addition, the correlation between extrovert, sensing, and feeling personality types and poetry compression were negative, but not significant.



Poems as literary materials can make learners more sensitive, self-conscious, and being aware of the world around (Carter & Long, 1991). For the reason that the judging learners are serious and meticulous (Oxford, 2003), their characteristics might help them in the process of poem comprehension. Moreover, the results of the present study confirm Strong's (1983) investigation on the influence of seven personality types on English proficiency. He found out that extraversion/introversion dichotomy of personality did not affect English proficiency. There is also no significant relationship between extraversion/introversion dichotomy and English poetry comprehension in the current study. Improved English poetry comprehension can be seen as a direct outcome of improved English proficiency. Hence, it is implied that extraversion/introversion dichotomy do not affect English poetry comprehension in the vein of the results of Strong's (1983) study. Also, the perceiving personality and L2 learners' poetry comprehension were negatively correlated in the present study. However, Pfister (2000, cited in Sadeghi & Mohd Kasim, 2012), using MBTI model of personality, found out that perceiving students' scores were better on interpretive comprehension items.

To embark upon the third research question concerning the best predictor of the poetry scores, a standard multiple regression analysis was carried out. Preliminary analyses were also conducted to ensure no violation of the assumptions of normality, linearity, multicollinearity, and homoscedasticity. To determine the more powerful predictor, scores on the personality types and gender were taken as predictor (independent) variables, and scores on the poetry were considered as the criterion (dependent) variable. The results are reported in Tables 4 and 5.

**Table 3.**

Model summary of multiple regression

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.087	.008	-.035	3.006

**Table 4.**

Results for regression analysis of personality types and gender

Model		Unstandardized Coefficients		Standardized Coefficients	<i>t</i>	Sig.
		B	Std. Error	Beta		
1	(Constant)	8.108	1.666		4.866	.000
	Gender	-.021	.941	-.003	-.022	.982



	Personality type	-.043	.074	-.087	-.589	.559
--	------------------	-------	------	-------	-------	------

As displayed in Table 3, the R Square was .008, indicating that the model which includes personality type and gender explains 0.8 percent of the variance in poetry scores, which is so small. As demonstrated in Table 4, none of the *t*-values for the coefficients were significant; thus, neither L2 participants' personality type nor their gender could make statistically significant contribution to the L2 learners' poetry comprehension. The findings, thus, suggest that gender and personality type cannot be a factor affecting the EFL learners' poetry comprehension.

Finally, an independent sample *t*-test was conducted to observe whether there was any significant difference between the male and female participants' poetry comprehension. The results of *t*-test on the poetry scores are shown in Table 5.

**Table 5.**

T-test on the male and female learners' poetry scores

	Levene's Test for Equality of Variances		T test for Equality of Means		
	<i>F</i>	Sig.	<i>t</i>	<i>df</i>	Sig.
Equal variances assumed	.903	.347	.124	48	.902
Equal variances not assumed			.130	29.827	.897

The male participants ( $M = 7.80$ ,  $SD = 2.731$ ) received higher poetry mean score than female participants ( $M = 7.69$ ,  $SD = 3.085$ ). As Table 5 demonstrates, the variance of the differences in the mean scores, according to the Levene's test, was equal ( $p = .347$ ), hence the equal variances were assumed for the *t*-test analysis. According to Table 5, there was no significant difference in poetry scores between males and females; the *t* value was so small,  $t(48) = .124$ ,  $p = .902$ . That is, the magnitude of the differences in the means was very small. The findings, thus, suggest that both male and female EFL learners can enjoy poetry comprehension and neither one can be advantageous on the level of poetry comprehension.



## 5. Discussion

Knowing personality profile of learners might help L2 teachers bring about new required methods and tasks. As Moradan (2006) states, poetry dealing with thoughts, things, feelings, and events which are the writers' real or imaginative experiences might help teachers personalize the learning process according to the learners' personality profiles. However, the results of this study did not support a strong positive correlation between the personality types and poetry comprehension among L2 learners. Even, some personality types were negatively correlated with the poetry comprehension. Based on the findings, there can be a significant relationship between judging/perceiving personality traits and understanding poetry. Language learners with judging personality trait can be advantageous on poetry comprehension. Thus, as Ehrman (2008) have pointed out, teachers should encourage L2 learners to build up their perception through guessing and taking out the intention of context and their judging behavior through scheduling and putting their studies and activities in order. This study, assumed to be a step in the right direction. Considering the limitations of the present study, further research is indeed needed to examine whether similar results can be gained. Also, the generalization of the findings should await other related research with larger samples in different L2 contexts.

## References

- Ahmazadeh, S. (2005). The nature and function of poetry. *Human Sciences*, 44, 1-10.
- Amilin Razawi, N., Muslim, M., Che Razali, S. M., Husin, N., & Abdul Samad, N. Z. (2011). Students' diverse learning styles in learning English as a second language. *International Journal of Business and Social Science*, 2(19), 179-186.
- Arthur, B. (1968). Reading literature and learning a second language. *Language Learning*, 18(3-4), 199-210.
- Ashton-Hay, S. (2004, March). *Teaching English poetry to Turkish undergraduates: Comprehension strategies matter*. Paper presented at Bilkent University Ninth International English Language Teaching Conference on Challenge in Learning: Helping Learners Realise Their Full Potential, Ankara, Turkey.
- Bachman, L. (1990). *Fundamental considerations in language testing*. Oxford: Oxford University Press.
- Braden, J. P. (1995). Intelligence and personality in school and educational psychology. In D. H. Saklofske & M. Zeidner (Eds.), *International handbook of personality and intelligence* (pp. 621-650). New York: Plenum.
- Callahan, S. (2000). Responding to the invisible student. *Assessing Writing*, 7, 57-77.



- Carrell, P. L. (1995). The effects of writers' personalities and raters' personalities on the holistic evaluation of writing. *Assessing Writing*, 2, 153-190.
- Carter, R., & Long, M. N. (1991). *Teaching literature*. London: Longman.
- Connell, R. W. (2002). *Gender*. Cambridge, England: Polity Press.
- DiTiberio, J. K., & Jensen, G. H. (1995). *Writing and personality: Finding your voice, your style, your way*. Palo Alto, CA: Davies-Black.
- Dubash, C. J., & Anwar, M. B. (2011). The integration of literature in second language acquisition and learning. *British Journal of Arts and Social Sciences*, 1(1), 35-40.
- Ehrman, M. E. (1994). The type differentiation indicator and adult foreign language learning success. *Journal of Psychological Type*, 30(1), 10-29.
- Ehrman, M. E. (2008). Personality and good language learners. In C. Griffiths (Ed.), *Lessons from good language learners* (pp. 61-72). Cambridge: Cambridge University Press.
- Ehrman, M. E., & Oxford, R. (1990). Adult language learning styles and strategies in an intensive training setting. *Modern Language Journal*, 74(3), 311-327.
- Finch, A. (2003). Using poems to teach English. *English Language Teaching* 15(2), 29-45.
- Greenberg, M. T., Weissberg, R. P., O'Brien, M. U., Zins, J. E., Fredericks, L., Resnik, H., & Elias, M. J. (2003). Enhancing school based prevention and youth development through coordinated social, emotional, and academic learning. *American Psychologist*, 58, 466-474.
- Hayakawa, S. I., & Hayakawa, A. R. (1990). *Language in thought and action* (5th ed.). New York: Harcourt, Brace, Jovanovich.
- Hess, N. (2003). Real language through poetry: A formula for meaning making. *ELT Journal*, 57(1), 19-25.
- Higgins, D. M., Peterson, J. B., & Pihl, R. O. (2007). Prefrontal cognitive ability, intelligence, Big Five personality, and the prediction of advanced academic and workplace performance. *Journal of Personality and Social Psychology*, 93, 298-319.
- Horning, A., & Sudol, R. (Eds.). (1997). *Understanding literacy: Personality preference in historical and psycholinguistic contexts*. Cresskill, NJ: Hampton Press.
- Jung, C. G. (1921). Psychological Types. In J. Campbell (Ed.), *The portable Jung* (pp. 178-269). New York: Viking Press.
- Jung, C. G. (1971). *Psychological types*. Princeton, New Jersey: Princeton University Press.
- Kiany, G. R. (1998). English proficiency and academic achievement in relation to extraversion-introversion: A preliminary study. *INJAL*, 8(1), 113-130.
- Koo, Y. L. (1992). *Selecting and exploiting printed authentic materials in the ESL context for the reading classroom: A Malaysian educator's experience*. In J. M. Ali (Ed.), *Proceedings on literacy in Asian societies*. Kuala Lumpur: Malaysian Reading Association.
- Lawrence, D. H. (1885-1930). City life. In L. Perrine (Ed.), *Literature: Structure, sound, and sense* (pp. 283-284). New York: Harcourt Brace Jovanovich.



- Matthews, G., Zeidner, M., & Roberts, R. D. (2006). Models of personality and affect for education: A review and synthesis. In P. A. Alexander & P. H. Winne (Eds.), *Handbook of educational psychology* (2nd ed., pp. 163-186). Mahwah, New Jersey: Lawrence Erlbaum Associates.
- Mayer, J. D. (2007). *Personality: A systems approach*. Boston: Allyn & Bacon.
- McPherson, B. (1999). Correlating students' personality types with their rating of topics covered in business communication classes. *Business Communication Quarterly*, 62(3), 46-53.
- Moradan, A. (2006). The miraculous power of poetry as a language teaching device in EFL classes. *TELL*, 1(1), 105-123.
- Myers, I. B., & McCaulley, M. H. (1985). *Manual: A guide to the development and use of the Myers-Briggs Type Indicator*. Palo Alto, CA: Consulting Psychologists Press.
- Myers, I. B., McCaulley, M. H., Quenk, N. L., & Hammer, A. L. (2003). *Manual: A guide to the development and use of the Myers-Briggs type indicator* (3<sup>rd</sup> ed.). Palo Alto, CA: Consulting Psychologists Press.
- Myers, I. B., & Myers, P. B. (1980, 1995). *Gifts differing: Understanding personality type*. Mountain View, CA: Davies-Black Publishing.
- Noftle, E. E., & Robins, R. W. (2007). Personality predictors of academic outcomes: Big five correlates of GPA and SAT scores. *Journal of Personality and Social Psychology*, 93, 116-130.
- Oxford, R. L. (2003). Language learning styles and strategies: An overview. *Learning Styles & Strategies*, 1, 1-25.
- Padmanugraha, A. S. (2011, May). *Poetry, gender and teaching: Building students' character through multicultural literature*. Paper presented at the Third Conference on Teaching English as A Foreign Language (COTEFL) 'Multiculturalism, Media, and the Modes of Best Practices,' Muhammadiyah University of Purwokerto, Indonesia.
- Piske, T., MacKay, I. R. A., & Flege, J. E. (2001). Factors affecting degree of foreign accent in an L2: A review. *Journal of Phonetics*, 29, 191-215.
- Pittenger, D. J. (1993). Measuring the MBTI ... and coming up short. *Journal of Career Planning and Employment*, 54, 48-53.
- Rahimi, M. (2005). *An investigation into the factors affecting the use of language learning strategies by Persian EFL learners*. Unpublished Doctoral dissertation, Shiraz University, Iran.
- Rajendra, N. M. (1992). *Search for belief in the poetry of Robert Frost*. India: Abhinav Publications.
- Sadeghi, N., & Mohd Kasim, Z. (2012). Learning styles, personality types and reading comprehension performance. *English Language Teaching*, 5(4), 116-123.
- Sharp, A. (2008). Personality and second language learning. *Asian Social Science*, 4(11), 17-25.
- Shehade, A. (1999). Gender differences and equal opportunities in ESL classrooms. *ELT Journal*, 53(4), 255-261.



- Sidney, P. (1554-1586). *With how sad steps, o moon, thou climb'st the skies.* Retrieved September 12, 2011, from the World Wide Web: <http://rpo.library.utoronto.ca/poem/1930.html>
- Stake, J. E., & Eisele, H. (2010). Gender and Personality. In J. C. Chrisler, & D. R. McCreary (Eds.), *Handbook of gender research in psychology* (pp. 19-40). New York: Springer Press.
- Strong, M. H. (1983). Social styles and second language acquisition of Spanish speaking kindergarteners. *TESOL Quarterly*, 17(2), 241-258.
- Sundberg, N. D. (1970). Critique of the Myers-Briggs type indicator. In O. K. Buros (Ed.), *The sixth mental measurement yearbook* (pp. 1127-1130). Highland Park, Gryphon.
- Teika, F. N. (13<sup>th</sup>). A travel poem. In E. Minter (Ed.), *An introduction to Japanese court poetry* (p. 115). Stanford: Stanford University Press.
- Teika, F. N. (13<sup>th</sup>). A winter poem. In E. Minter (Ed.), *An introduction to Japanese court poetry* (p. 115). Stanford: Stanford University Press.
- Thompson, T. C. (1995). Understanding attitudes toward assessment: The personality factor. *Assessing Writing*, 2, 191-206. Under the el-train. (n.d.). [http://www.ncjrs.gov/html/ojjdp/news\\_acglance/216684/topstory.html](http://www.ncjrs.gov/html/ojjdp/news_acglance/216684/topstory.html)
- Wakamoto, N. (2007). *The impact of extroversion/introversion and associated learner strategies on English language comprehension in a Japanese EFL setting*. Unpublished doctoral dissertation, University of Toronto, Canada.
- Yung, Y. (n.d.). The river. In V. Baron (Ed.), *Sunset in a spider web: Sijo poetry of ancient Korea* (p. 34). New York: Holt, Rinehart & Winston.

### Bio Data

**Ali Roohani** is a PhD holder in TEFL. He received his BA in English literature, his MA and PhD in TEFL from Shiraz and Allameh Tabatabai' universities, respectively. He is currently a faculty member at Shahrekord University. His area of interest includes EQ, multiple intelligences, learning strategies and motivation. He has published 3 books and 12 papers and presented a number of papers in the national and international conferences.

**Faranak Forouzandeh** got her B.A. in English Translation from Shahrekord University and is an M.A. TEFL graduate of the same university. Her areas of interest include literature, psycholinguistics, sociolinguistics, and cognitive linguistics.



# English Language Teachers' Perceptions of Educational Supervision in Relation to Their Professional Development: A Case Study of Iran

Zargham Ghapanchi<sup>1</sup>

Marjan Baradaran

*Ferdowsi University of Mashhad*

## Abstract

This study indicates the nature of instructional supervision carried out in SAFIR English Language Academy and Ferdowsi University of Mashhad- Iran. It followed the qualitative method involving questionnaire and interview. The paper commences with theoretical underpinnings of dual terms, educational supervision and professional development in relation to English language context. 45 teachers participated in this study and also interviews were conducted face to face. The results of the study show that educational supervisors highlight the strengths of English language teachers' performances and encourage them for further practice to remove their weaknesses and overcome them. However, some problems stated by the teachers participated in this survey regarding educational supervision which will be pointed out later in detail. Two important aspects of language teaching were discussed and concluded in this study which are on one side the remarkable difference between language teachers' belief and real experiences concerning supervision and on the other side their expectations of effective supervision by supervisors.

**Keywords:** Educational supervision, Supervision, Professional development

## 1. Introduction

There is a growing consensus in the literature regarding the elements of effective professional development for teachers. It incorporates principles of adult learning: Adult learners need to be self-directed; they display readiness to learn when they have a perceived need; and they desire immediate application of new skills and knowledge (Knowles, 1980). Effective professional

---

<sup>1</sup> Email: ghabanchi@um.ac.ir



development is embedded in the reality of schools and teacher s' work. It is designed with teacher input. It fosters critical reflection and meaningful collaboration. It is internally coherent and rigorous, and it is sustained over the long term (Little, 1993; Renyi, 1996; Sparks & Hirsch, 1997). Promising professional development is organized with both effective teaching and learning.

These elements underlie various professional development structures such as university-school partnerships (Darling-Hammond, 1997), teacher networks and collaboratives (Little, 1993; Renyi, 1996) and teacher study groups (Clair, 1995; 1998), to name a few. What these structures have in common is opportunities for teachers to feel responsible for the professional development process to be knowledge creators as opposed to mere receivers of information.

To maintain ongoing professional development, English language teachers get involved in many professional activities or build up their own self-development strategies either individually or collaboratively. Also peer-coaching, study groups, action research, teaching portfolios, team teaching, and in-service training are some other effective professional development strategies. Cark (cited in Hargreaves and Fullan, 1992, p. 81) proposes that professional development is basically 'a solitary journey'; however, almost all teachers need assistance and support during that journey from colleagues or supervisors to enhance their own development, by which they can gain an inside perspective on other teachers' experiences and raise their awareness via reflecting on their own situation. At this point, Edge (cited in Bailey et. al., 2001) notes that working in isolation holds teachers back and subjective experience shared with no one cannot contribute to their development, but through the quality collaboration teachers have a chance to eliminate subjectivity and draw some conclusions regarding their experiences and opinions.

Many studies were conducted on educational supervision in teachers' professional development; for instance, Wallace (1991), Ur (1996), Freeman and Johnson (1998), Freeman (2001), and Richards and Farrell (2005) write about teachers' professional development by emphasizing the vitality of self-development in their career. Moreover, Okeoforth and Poole (1992) studied on the role of educational supervision in teachers' professional growth. In this vein, educational supervision, a cooperative problem solving process, can be regarded as a key feature in English language teachers' professional development.

However, there is still no agreement on the precise definition of the term 'educational supervision' since some differences have been mentioned related to orientations, perceptions, comprehension and familiarity with aspects of the framework and also analysis of its content in different studies. To illustrate that point, Allan (1990) describes 'educational supervision' as a set of duties and a comprehensive process which aim to help teachers to develop their profession to achieve their



pedagogical objectives while Daresh (2001, p. 25) defines this term in thorough way as a dynamic process that leads to studying and improving all factors that affect the education situation. On the other hand, according to Kilminster et al. (2007, p. 2), educational supervision is a matter of guidance and feedback regarding personal, professional and educational development in the context of trainee's experience. Bailey's (2006) point of view, the modern educational supervision is characterized as follows is a technical process which aims to improve teaching and learning through the care, guidance and simulation of continued development for not only teachers but also any other person having an impact on the educational context.

It is a consultation process, based on respect for the opinion of teachers who are mainly affected by the work of supervision. It is a collaborative process in different stages since it welcomes various views that represent the proper relationship between the supervisor and the teacher so as to address the educational problems and find appropriate solutions.

It is an academic process which encourages research and experimentation whose results can be used to improve setting and achieving clear, observable and measurable objectives in the educational setting.

It is a leadership process which requires the supervisor to have the ability to coordinate teachers' efforts by aiming to achieve the teaching objectives.

It is a humanitarian process in which the supervisors recognize the value of individuals as human being so that they can build a mutual trust between themselves and the teachers and know the exact and varying capacities of each teacher they deal with.

By considering all these definitions, some functions are worth mentioning here (Brown & Bourne, 1995; Kadushin, 1992; Salaman, 1995; Smith, 2005), which are grouped into six main categories stated as follows:

1. Development of curricula
2. Supervision and organization of the educational setting
3. Supervision of the teachers' professional development
4. Supervision of the teaching methods and techniques
5. Supervision of the novice teachers
6. Evaluation of the educational process



The issues surrounding supervision have been debated by teachers, administrators, higher education scholars, and legislators (Gland & Neville, 1997; Glickman, 1992; Sergiovanni & Starratt, 2002).

According to Glickman (1992), "without a strong, effective, and adequately staffed program of supervision, an effective school is unlikely to result". Supervision can be defined as "the glue of a successful school" (Glickman, Gordon, & Ross-Gordon, 2007). Supervision has become a major component in the operation of schools (Sergiovanni & Starratt, 2002).

### *1.1. Supervision as a tool for improvement*

Goldhammer (1969) suggested that supervision should increase teachers' willingness and ability to supervise themselves and their colleagues. Zepeda (2003) advocated the term, "auto supervision," to describe the ability of teachers to supervise themselves, and has provided the sense for peer coaching. Supervision can be seen as similar to teaching in that teachers wish to improve students' behavior, achievement, and attitudes while supervisors wish to improve teachers' behavior, achievement and attitudes (Glickman, 1992).

### *1.2. Supervision as an obstacle*

Zepeda and Ponticell (1998) conducted a study to determine what teachers "need, want, and get from supervision". Their findings supported the positional power of the supervisor dominating over the inferior teacher. Blumberg (1980), in *Supervision and Teachers*: described the negative relationship between supervisors and teachers, describing the resentment teachers felt toward supervisors, and this resentment continues to be a major barrier in achieving benefit from the practice of supervision.

### *1.3. Objectives of the study*

The study aims to answer the following questions

- 1) What is the nature of supervision in English language institutes?
- 2) What are the perceptions of the experienced and inexperienced teachers regarding supervision practiced by the supervisors?
- 3) Are the teachers really benefited by the process of supervision?



## 2. Method

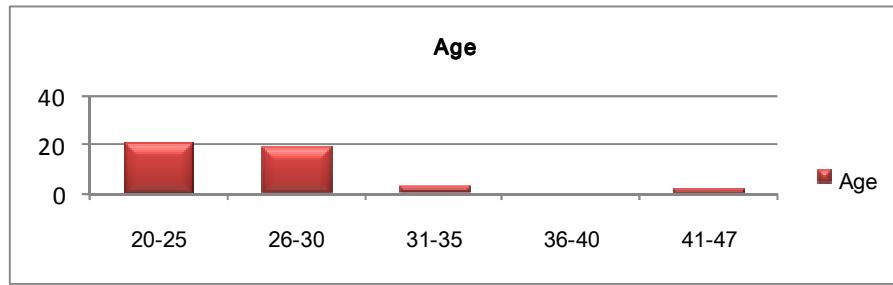
The aim of this study is to find out through qualitative and quantitative analysis of collected data whether the educational supervision plays a minor or major role in English language teachers' professional development and also what language teachers' perceptions are concerning educational supervision.

### *2.1. Participants*

Forty five non-native teachers, thirty of them working at Safir English Language Academy in Mashhad-Iran and fifteen of them were Ferdowsi University students (M.A.) which were teaching English in different schools or institutes participated in this study. Teachers' ages ranged from 20-47 years-old with an average age of 27.

The number of female teachers was higher than that of male teachers as seen in the graphic below (Figure 1). The educational system in Safir English Language Academy consists of three types of classes: Termic, Intensive and Fridays which overall provides twenty sessions each semester. Language teachers in this institute have different qualifications such as: FCE, CAE, CPE, IELTS certificates and most of them are university graduates .When they pass the first step which requires the head office (Tehran) confirmation, they will have a week training orientation. If they meet the requirements, they observe the experienced teachers classes in order to be familiarized with the system and gain some experiences.

Furthermore, in this system the teachers are observed once or more each semester and they receive some feedback in a session called "briefing session" relevant to other activities that will support teachers professionally. It can be pointed out that teacher meetings, on job training programs (OJT) and teacher training courses (TTC) are also organized to which attendance is of great importance.



**Figure 1.**

The number of male and female teachers with respect to their age

## 2.2. Research instruments

### 2.2.1. Pre-questionnaire

A pre-questionnaire was administered to the teachers before they started answering the post-questionnaire. This section asked the subjects their age, gender, previous experience in the related field.

### 2.2.2. Post-questionnaire

A post-questionnaire was administered to the teachers after their last English class session in 2010-2011 Academic Year. The purpose of the post-questionnaire was to find out English language teachers' perceptions of educational supervision in relation to their professional development which was studied on these seven subcategories:

1. Teachers' experience
2. Educational leadership and humane relations
3. The curriculum and teaching methods / techniques
4. The students
5. Teaching materials and activities
6. Classroom management
7. Assessment

The content validity of this questionnaire was established through a review of five professors working in universities, and then some slight changes were made based on their recommendations. Later, the questionnaire was piloted in a group of 12 English language teachers, who were also



asked to comments on the appropriateness and ease of the survey instrument, which led to simplifying the wording of some items. After all revisions, to make sure of the reliability, the Cronbach's Alpha Coefficient was used and the reliability was (0.88) for the questionnaire.

#### *2.2.3. Interviews*

The researcher with the aid of the supervisor interviewed with each other and the teachers to cross-check teachers' responses to the post-questionnaire and to find out more information on their responses which was not possible in the post-questionnaire. Hence, interview questions contained three items:

1. Do you give enough importance to your professional development? If no, what impedes you from keeping your continuous professional development?
2. Do you think that educational supervision is significant for your professional development?
3. Which ways of educational supervision are you familiar with?

#### *2.3. Data analysis*

Fifty copies of the questionnaire were distributed to the targeted sample at the end of 2010-2011 Academi Year to choose the suitable statement given five point Likert-scale, ranging from Strongly Agree (SA) to Strongly Disagree (SD) and the data was tabulated and analyzed by using the Statistical Package for Social Sciences (SPSS) and Microsoft Office Excel.

### **3. Results**

#### *3.1. Post-questionnaire*

As can be seen from Table 1, the majority of responses for most of statements are on the agree side, which implies that most of the teachers contend that educational supervision positively effects on developing their experiences in their field. Moreover, when looking specifically at the second item, it appears that 62% of the teachers strongly believe that supervisors encourage the teachers to be aware of the latest issues in the field. On the other hand, for statement 6 and 7, the number of neutral responses are more than other statements, which shows the lack of case studies and also post-graduate studies among teachers. Also, they perceive the supervisors as a professional who specifies teachers' needs and encourages them to update their educational experiences.



**Table 1.**

Teachers' perceptions about the effect of educational supervision on English language teachers' professional development in terms of developing experiences in the field

No	The main fields and their sections	SA	A	N	D	SD
	1) Developing the teacher's experience					
1	Prepare the teacher's mind to accept the development by explaining the importance of the latest educational issues.	0.56	0.37	0.05	0.02	
2	Encouraging teachers to be aware of the latest educational issues in order to exchange successful experience with their colleagues.	0.66	0.32	0.02	0.00	
3	Guide the teachers to the books, research papers, and educational journals which may help them with professional development.	0.32	0.54	0.07	0.05	
4	Hold meetings with the teachers to study and discuss the educational circulars sent by Ministry of Education.	0.29	0.54	0.07	0.07	
5	Encourage the teachers to the self- learning to update profession experiences.	0.56	0.39	0.05	0.00	
6	Encourage the teacher to conduct case- studies to solve some problems that they encounter.	0.32	0.32	0.29	0.05	0.02
7	Encourage the teachers to continue their post-graduate studies to raise their professional and academic levels.	0.29	0.37	0.20	0.12	0.02
8	Direct the teachers to benefit from TV and radio computer programs and the Internet.	0.39	0.46	0.10	0.05	
9	Participate in organizing educational workshops, conferences, and symposiums and invite specialists in education to give lectures.	0.46	0.41	0.05	0.02	0.02
10	Participate in organizing educational workshops, conferences, and symposiums and encourage teachers to give lectures.	0.29	0.49	0.17	0.02	0.02
11	Nominate the teachers to attend educational workshops to encourage them to develop professionally.	0.37	0.49	0.12	0.02	



12	Specify the teachers' professional development needs.	0.51	0.41	0.05		
13	Advise the teachers to participate in different courses which help them to update their educational experiences.	0.54	0.37	0.10		

**Table 2.**

Teachers' perceptions about the effect of educational supervision on English language teachers' professional development in terms of educational leadership and human relations.

No	The main fields and their sections	SA	A	N	D	SD
	2) Educational leadership and humans relations					
14	Allows the teachers to act freely towards their duties to be accustomed to carrying out responsibilities.	0.34	0.37	0.15	0.15	
15	Distributing responsibilities among the teachers in a suitable way relevant to their knowledge and professional abilities.	0.20	0.56	0.17	0.02	
16	Give opportunity to the teachers to express their professional opinions to gain benefits from them.	0.46	0.44	0.05	0.02	

Table 2 suggests that almost all teachers agree on the positive effects of educational supervision in terms of educational leadership and human relations. However, for statement 14 there are some answers on disagree side and neutral side. Although the percentage is low (15%), it seems that some problems occur as to the case that some systems do not let teachers act freely.

**Table 3.**

Teachers' perceptions about the effect of educational supervision on English language teachers' professional development in terms of curriculum and teaching methods.

NO	The main fields and their sections	SD	A	N	D	SD
	3) The curriculum and teaching methods/techniques					
17	Train the teachers to use modern teaching techniques.	0.44	0.49	0.02	0.02	



18	Train the teachers to develop students' thinking skills.	0.59	0.20	0.02	0.02	
19	Train the teachers to develop students' communication skills.	0.63	0.29		0.02	
20	Direct the teachers to learn about the results of educational researches done on teaching techniques to use the most appropriate ones in their teaching.	0.24	0.54	0.15	0.05	
21	Give opportunity to the teachers to use suitable teaching techniques for their students.	0.44	0.49	0.05		
22	Conduct workshops, seminars, or micro-teaching to the English teachers in the schools.	0.37	0.49	0.12	0.02	
23	Make the teachers familiar with the latest updated issues about the curriculum.	0.46	0.37	0.07	0.05	
24	Involve the teachers, in the curriculum development process.	0.32	0.39	0.12	0.17	
25	Help the teachers with preparing their teaching plans.	0.46	0.44	0.02	0.07	
26	Let the teachers discuss ways of solving any problems with the curriculum.	0.41	0.49	0.07	0.02	

It is very vivid that in Table 3 nearly all responses are on the agree side by indicating the belief on positive impact of educational supervisor in terms of the curriculum and teaching methods and techniques. Considering statement 24 some responses appear on neutral side and disagree side which implies that teachers approve of being involved in curriculum development.

**Table 4.**

Teachers' perceptions about the effect of educational supervision on English language teachers' professional development in terms of the students.

No	The main fields and their sections	SA	A	N	D	SD
	4) The students					
27	Guide teachers to identify the students' developing characteristics to create positive learner attitudes towards learning.	0.44	0.44	0.05	0.07	



28	Direct teachers to develop their skills in dealing with the students' varied behaviors.	0.51	0.39	0.02	0.07	
29	Guide teachers to use effective methods to help students to overcome the difficulties that they face in learning.	0.44	0.44	0.12		
30	Help teachers how to deal with outstanding students.	0.37	0.49	0.07	0.05	0.02
31	Discuss the ways of treating the learners' weaknesses.	0.44	0.49	0.05	0.02	

The same continues here. What is put forward here is that the teachers really expect the supervisors to collaborate with them in order to solve the problems related to the students through discussions, which will help their professional development to a great extent. What is striking here is that there isn't any strongly disagree response in the first three items which proves our previous perceptions. Also, for statement 30, there are some responses on neutral and disagree side which implies that the dealing with outstanding students is problematic for teachers.

**Table 5.**

Teachers' perceptions about the effect of educational supervision on English language teachers' professional development in terms of teaching materials and activities.

No	The main fields and their sections	SA	A	N	D	SD
	5) Teaching materials and activities					
33	Direct teachers to use classroom activities which develop the learners' thinking skills.	0.41	0.54	0.02	0.02	
34	Encourage the teachers to carry out extra-activities (e.g. English club, magazines, competitions...)	0.41	0.39	0.15	0.05	

Table 5 indicates that 95% for statement thirty-three and 80% for statement thirty-four of the population agree on the positive role of supervisors in teachers' professional development. They believe that the supervisors encourage them to use extra and new classroom activities and carry out some activities thereby the students will receive novel information and develop their thinking skills.



**Table 6.**

Teachers' perceptions about the effect of educational supervision on English language teachers' professional development in terms of classroom management.

No	The main field and their sections	SA	A	N	D	SD
	6) Classroom management					
35	Make the teachers familiar with different techniques of classroom management.	0.54	0.41		0.02	
36	Guide the teachers in encouraging their students to express their ideas.	0.54	0.44	0.02		
37	Direct teachers to establish a positive relationship with the students.	0.56	0.27	0.05		
38	Explain to the teachers the ways of considering and teaching students with mixed abilities.	0.46	0.41	0.07	0.05	

As can be seen from Table 6, the supervisors are expected to hold meetings so that teachers can learn recent ways of classroom management and reflect upon their own practices. Also, guide the teachers to encourage their students to express their ideas. Moreover, supervisors are desired to direct teachers to establish a positive relationship with the students and endure the active participation of the learners during the lesson which is agreed by the majority of the teachers.

**Table 7.**

Teachers' perceptions about the effect of educational supervision on English language teachers' professional development in terms of assessment.

No	The main field and their sections	SA	A	N	D	SD
	7) Assessment					
39	Make the teachers familiar with assessment techniques (i.e. continuous assessment, formal testing...).	0.32	0.41	0.02	0.05	
40	Encourage the teachers to take notes about students' in-class performance and use them as assessment record.	0.39	0.41	0.1	0.1	
41	Train the teachers to design effective exams to be used in their classes.	0.22	0.39	0.27	0.12	



42	Provide the teachers with recent assessment techniques.	0.32	0.46	0.1	0.1	
43	Train the teachers to ask the students variety of oral questions that are more suitable to the teaching-learning situation.	0.39	0.49	0.05	0.07	
44	Guide the teachers to use practical ways of diagnosing the strengths and weaknesses of students.	0.46	0.49		0.05	
45	Encourage teachers to benefit from the assessment results to suggest ways and solutions for the difficulties that the students face.	0.39	0.54	0.02	0.05	

Table 7 shows that the great number of the teachers agreed that supervisors are to update the teachers' knowledge about the techniques used for assessment since there is continuous change in every educational system and also the supervisors are expected to help the teachers diagnose the strengths and weaknesses of test types and modify them according to teaching and learning situation. Furthermore, as the item forty-five suggest, the supervisors should make recommendations about how to use the assessment results to suggest solutions for all challenges.

### 3.2. Interviews

The researchers interviewed with each teacher to reveal their specific ideas regarding the effect of educational supervision on their professional development. As stated earlier, interview questions contained three items. Interview questions were posed to each teacher. The researchers transcribed language teachers' responses to the questions as they heard from their utterances.

In response to Question 1 regarding whether they give enough importance to their professional development or not, 25 (62%) teachers responded positively; yet, 16 (40%) expressed that they are not able to think of their professional lives due to many reasons in both their personal and professional lives. To illustrate,

*... Motivation, devotion and love are the factors you need to have in your job in order to obtain what is prepared for you. As a teacher I have rarely been prepared by feedback on my work except SAFIR INSTITUTE also I believe you need to be fed by the system to be able to work and think of professional development which I got both in my work.*



*...I might give enough importance to it but I feel like sometimes I'm not quite aware of all the sources available and plus working sometimes doesn't leave me enough time to develop as much as I want.*

*...Frankly I didn't give enough importance to my professional development but recently I' trying to get rid of unnecessary issues and focus on my job.*

*...actually I give enough importance to my professional development but having too many classes and writing the lesson plans doesn't leave me enough time to do anything else.*

In relation to Question 2 regarding the significance of educational supervision in their professional lives, teachers displayed controversial ideas, which reflect the diversity between what they believe in and what they actually experience. Almost all of the teachers noted that they consider educational supervision as a concept that can contribute to their professional progress which is expressed as follows:

*...I think that educational supervision has been significant for my professional development but sometimes supervisors can be too critical and pay too much attention to unimportant details that only confuses the teacher and makes the teacher feel insecure about every activity he or she designs.*

*...A third eye usually detects weaknesses and problems better.*

*... Under some conditions I agree with the idea of educational supervision:*

*...Continuous educational supervision shouldn't lead to an atmosphere of anxiety and stress so that the teachers don't feel emotionally drained.*

*...The observer should be highly trained and able to maintain an objective point of view.*

*...definitely yes especially for me that I'm majoring in an irrelevant field, so using the supervisors experiences will be of great help.*

As to Question 3 related to which methods of educational supervision they are familiar with, they mostly weren't aware of all educational supervisions and they just mentioned class observations and peer observation which those who are working in SAFIR institute are quite familiar with, since teachers are observed in this institute at least once per class. For instance some noted that:



*...class and peer observations which we are quite familiar with these methods in our institute and are very helpful.*

*...seminars and workshops and also teacher's meetings are considered as educational supervision.*

#### **4. Discussion**

The findings of the research reveal that English language teachers seem to show a very positive attitude towards professional development whereby they are able to build on previous activities by learning new trends in the field, discuss their experiences with other teachers, reflect upon their own teaching, and gain practical experiences of other opportunities encouraging their ongoing professional development, which is the main point in the recent educational system that accepts the significance of effective and continuous professional development (CPD) (Birman et al., 2000; Borko, 1995; Ingvarson et al., 2005; Lowden, 2006). However, there are major impediments on the way of teachers' desired professional development, which are allocated time and financial resources. In particular, the teachers need time both to make professional development a continuous part of their work (Bush, 1999) and to see the outcomes of their efforts (Dorph & Holtz, 2000), which clearly indicate that lack of time is one of the greatest challenges to implementing effective professional development (Abdal-Haqq, 1996). As to the financial resources, although funding is of great importance in each educational organization or effective professional development, it is usually quite low even though its significance in teachers' professional development was reported in many studies conducted all around the world.

However, as Bush (1999) points out, overcoming these impediments or problems is not enough to provide effective professional development opportunities for teachers. What is needed by many teachers is ongoing support and guidance (Lind, 2007). To put it differently, the main professional need for many teachers is to obtain practical ideas, strategies, and resources that can be applied in or outside the classroom (Beale, 2003). It is the educational supervisor who can assist the teachers to develop themselves professionally. Similarly, several studies report the leadership of supervisors is of utmost importance to support teachers' professional development in schools (Fernandez, 2000; Moore, 2000; Clement & Vandenberghe, 2001; Marshall et al., 2001). As for the characteristics of a supervisor, Robbins (1999) stresses that a supervisor is to provide the teachers with support, guidance, feedback, problem-solving skills, and a network of colleagues who share resources, insights, practices, and materials. These characteristics of a supervisor are totally in line with how English language teachers conceive educational supervision. However, certain misinterpretations have been noticed in the light of interviewed teachers' perspectives when going beyond their perceptions and considering their experiences in a virtual setting .In particular, the language



teachers believe that educational supervisors attempt to detect teachers' mistakes in classrooms, focusing mainly on control, as in the first stage of supervision lasting till 1860s (Daresh, 2001). Also, they regard the supervisors as people who look at their job performance in a judgmental way. This situation shows that educational supervision does not extend beyond the traditional notion of supervision that evokes the sense of forcing them to prescribe to a specific educational philosophy which will lead to a less desirable effect on the teaching process. Instead, educational supervision should provide the teachers with adequate support to encourage them to develop their own ways and means for using their own capabilities that will create a culture of collaboration or community of practice (Sergiovanni, 2005).

Furthermore, The teachers consider that some severe problems are involved in this control-based applications of educational supervision (Sullivan & Glanz, 2000) such as the method used for supervision and the way it is carried out.

However, such methods of educational supervision as individual interview, group meetings, exchange visits between teachers, model lesson, educational research, educational circulars, educational training, and so forth, can be applied according to the goal of educational improvement, requirements of the educational supervision, the surrounding environmental conditions, and the characteristics of the teachers (Alhabeeb, 2004; Bailey, 2006; Mone, 1998; Richards et al., 2001, 2005; Wallace, 1991). In fact, classroom visits can extremely be beneficial for the teacher on condition that they are conducted in an appropriate way; otherwise, some adverse effects may emerge on the teachers' classroom performance that can be from being observable, comprehensive and distinguishing in nature and there may be great differences between the results of supervisors' observations on the same teacher at different times, which can yield a decline in the teachers' trust and respect towards supervision (Memduhoglu et al., 2007). At this juncture, the huge importance of communication and collaboration becomes very clear, so the unpredictable nature of classroom visits or other methods of educational supervision needs certain changes. Before supervision takes place, supervisors are expected to have a brief conference with the teacher to be observed and inform her/him about the style and the form of the supervision, then enter the classroom together with the teacher, not during the lesson (Mone, 2001); however, this happens in an unexpected fashion in practice. Moreover, the supervisors are supposed to hold a meeting including the staff, during which observations and evaluations are shared and additionally, they think that a supervision report should be written about observations, defects, improper and desirable practices so as to make some suggestions and provide future insights for the coming supervision (Açıkgoz, 2001; Taymaz, 2001). In brief, supervision is to be regarded as a process of interpersonal relationships whereby people would listen to each other, bring alike problems, come together with



resources to find solutions and feel secure in the educational process (Wiles & Bondi, 1996), which is, according to Sergiovanni and Starratt (1993), very important in achieving school goals at school.

The perspectives discussed above briefly suggest that a great number of teachers are of the opinion that educational supervision brings about emergent outcomes to the overall teaching and learning community in case it provides a continuum of collaborative relationship where teaching experiences are reflected upon and shared with other staff in building a community of practice (Downey et al., 2004; Glickman et al., 2007). In this respect, the term, reflection, begins to appear as a vital phenomenon to be encouraged and promoted during the supervision process (Clarke, 1995) by which the teacher:

- becomes curious about certain aspect of practice setting;
- frames aspect in terms of the particulars of the setting;
- reframes that aspect in the light of past knowledge and previous experience;
- develops a plan for future action.

All in all, the findings which can be drawn from this study are on one side the great difference between English language teachers' beliefs and actual experiences about educational supervision and on the other side their expectations concerning effective applications of educational supervision to be managed by effective supervisors who are able to (Memduho et al., 2007):

- observe and reflect on practice
- give constructive feedback
- teach
- identify alternatives
- solve problems
- motivate
- foster autonomy
- provide information
- appraise self and others
- manage a service
- create a supportive climate
- negotiate
- manage time
- organize



The feedbacks from this study reveal that still observation is conducted on traditional basis in most educational systems. However, there are some institutes which are doing well in improving this trend. Just as the knowledge based for teaching foreign languages remains highly contested, what a supervisor needs to know and be able to do and where that knowledge comes from are all questions that remain little addressed in the research or practice literature for L2 teaching and learning. That said, based on our experiences working with teachers and supervisors, we have tried to make the case here that effective supervision requires flexibility and, perhaps most of all, attention to the bigger picture of sustainable teacher development. Lastly, it is essential to note that although important progress has been made in most educational supervision systems today, further reforms are necessary to renew the purpose, function, method and structure of educational supervision.

## References

- Abdal-Haqq, I. (1996). *Making time for teacher professional development* (Digest 95-4). Washington, DC: ERIC Clearinghouse on Teaching and Teacher Education.
- Alhabeeb, A. (2004). *Educational supervision ways in teacher's professional development*. AlReyadh: King Suod University.
- Allan, G. (1960). *Supervisory leadership*. London: Scott, Foresman.
- Applewhite, A. S. (1999). Funding staff development for school improvement and student achievement. *NASSP Bulletin*, 83(610), 49-54.
- Bailey, K. M. (2006). *Language teacher supervision: A case-based approach*. Cambridge: Cambridge University Press.
- Bailey, K. M., Curtis, A., & Nunan, D. (2001). *Pursuing professional development: The self as a source*. Boston, M.A: Heinle & Heinle.
- Beale, J. (2003). *What do English language teachers want from professional development?* Unpublished M.A thesis. Melbourne: Monash University.
- Birman, B. F., Desimone, L., Porter, A. C., & Garet, M. S. (2000). Designing professional development that works. *Educational Leadership*, 57(8), 28-33.
- Brown, A. & Bourne, I. (1995). *The social work supervisor: Supervision in community, day care and residential settings*. Buckingham: Open University Press.
- Bush, W. S. (1999). Not for sale: Why a bankroll alone cannot change teaching practice. *Journal of Staff Development*, 20(3), 61-64.
- Clarke, D., & Hollingsworth, H. (2002). Elaborating a model of teacher professional growth. *Teaching & Teacher Education*, 18(8), 947-967.
- Clarke, A. (1995). Professional development in practicum settings: reflective practice under scrutiny. *Teacher & Teacher Education*, 11(3), 243-261.



- Clement, M. & Vandenberghe, R. (2001). How schools leaders can promote teachers professional development: an account from the field. *School Leadership and Management, 21*(1), 43-57.
- Downey, C. J., & Frase L. E. (2004). *Participant's manual for conducting walk-through with reflective feedback to maximize student achievement* (2nd ed.). Huxley, IA: Curriculum Management Services.
- Fernandez, A. (2000). Leadership in an era of change: breaking down the barriers of the culture of teaching. In Day, C., Fernandez A., Hauge, T. E., Moller, J. (Eds.), *The life and work of teachers: International perspectives in changing times*. (pp. 239-255). London: Falmer Press.
- Freeman, D. (2001). Second language teacher education .In R. Carter & D. Nunan (Eds.). *The Cambridge guide to teaching English to speakers of other languages*. (pp. 72-79). Cambridge: Cambridge University Press.
- Freeman, D., & Johnson, K. E. (1998). Reconceptualizing the knowledge base of language teacher education. *TESOL Quarterly, 32*(3), 397-417.
- Fullan, M. (1991). *The new meaning of change*. New York: Teachers College Press.
- Glickman, D. C., Gordon, S. P., & Ross-Gordon, J. M. (2007). *Supervision and instructional leadership*. Pearson Education Inc.
- Hamilton, J. (1996). *Inspiring innovations in language teaching*. Clevedon: Multilingual Matters.
- Hancock, R. (2001). Why are class teachers reluctant to become researchers? In Soler, J., Craft, A., & Burgess, H. (Eds.), *Teacher development: exploring our own practice* (pp. 119-132). London: Paul Chapman Publishing and the Open University.
- Hargreaves, A., & Fullan, M. G. (1992). *Understanding teacher development*. London: Cassell.
- Haycock, K. (1999). The national library power project. *Teacher Librarian, 27*(2), 34. Retrieved from Academic Search Complete database.
- Hismanoglu, M., & Hismanoglu, S. (2010). English language teachers, perceptions of educational supervision in relation to their professional development, *Novitas Royal (Research on Youth And Language), 2010, 4*(1), 14-34.
- Ingvarson, L., Meiers, M., & Beavis, A. (2005). Factors affecting the impact of professional development programs on teachers' knowledge, practice, student outcomes & efficacy. *Education Policy Analysis Archives, 13*(10), 1-26.
- Karavas-Doukas, K. (1998). Evaluating the implementation of educational innovations: lessons from the past. In P. Rea-Dickins & K. P. Germaine (Eds.). *Managing evaluation and innovation in language teaching* (pp. 25-55). London: Longman.
- Kayıkçı, K. (2005). *Millî Eğitim Bakanlığı denetmenlerinin denetim alt sisteminin yapısal sorunlarına ilişkin algıları ve iş doyum düzeyleri*. Ankara: TEMSEN Yayıncıları.
- Lind, V. R. (2007). High quality professional development: An investigation of the supports for and barriers to professional development in arts education. *International Journal of Education & the Arts, 8*(2), 1-18.



- Lowden, K., McConnell, F., Hamilton, S., & Wilkinson, E. (2006). Evaluation of entrepreneurial spirit programme SEED. Edinburgh. Available online at: [www.flatprojects.org.uk/evaluations/evaluationreports/entrepreneurialspirit.asp](http://www.flatprojects.org.uk/evaluations/evaluationreports/entrepreneurialspirit.asp) (accessed 01 September 2009).
- Marshall, J. C., Pritchard, R. J., Gunderson, B. H. (2001). Professional development: what works and what doesn't? *Principal Leadership*, 1(6), 64-68.
- Memduoğlu, H. B., Aydin, I., Yilmaz, K., Güngör, S., & Oğuz, E. (2007). The process of supervision in the Turkish educational system: Purpose, structure, operation. *Asian Pacific Education Review*, 8(1), 56-70.
- Miller, J., McKenna, M., McKenna, B. (1998). A comparison of alternatively and traditionally prepared teachers. *Journal of Teacher Education*, 49(3), 165-179.
- Moore, K. B. (2000). Successful and effective professional development. *Early Childhood Today*, 15(3), 14-15.
- Okeafor, K. R., & Marybeth, G. P. (1992). Instructional supervision and the avoidance process. *Journal of Curriculum and Supervision*, 7(4), 372-392.
- Pachier, N., & Field, K. (1997). *Learning to teach modern languages in the secondary school*. London: Routledge.
- Richards, J. C., & Farrell, T. S. C. (2005). *Professional development for language teachers*. Cambridge: Cambridge University Press.
- Richards, J. C., Gallo, P. B., & Renandya, W. A. (2001). Exploring teachers' beliefs and the processes of change. *The PAC Journal*, 1(1), 41-62.
- Robb, L. (2000). *Redefining staff development*. Portsmouth: Heinemann.
- Robbins, J. (1999). *East meets west: Approaches to learner autonomy*. Proceedings of JALT 98, Japan Association for Language Teaching International Conference, Omiya, Japan.
- Salaman, G. (1995). *Managing*. Buckingham: Open University Press.
- Salas, S. & Mercado, L. (2010). *Looking for the big picture: macro strategies for L2 teacher observation and feedback*. United States and Peru
- Sandholtz, J. H. (2002). Interdisciplinary team teaching as a form of professional development. *Teacher Education Quarterly*, 27(3), 39-54.
- Sergiovanni, T. J., & Starratt, R. J. (1998). *Supervision: A redefinition*. New York: McGraw-Hill.
- Smith, G. (2000). Research and inspection: Her majesty's inspectorate of schools and the office for standards in education, 1981-1996 A commentary. *Oxford Review of Education*, 26(3), 333-352.
- Smith, B. (2005). Job retention in child welfare: Effects of perceived organizational support, supervisor support, and intrinsic job value. *Children and Youth Services Review*, 27, 153-169.



- Sullivan, S., & Glanz, J. (2000). *Supervision that improves teaching strategies and techniques*. California: Sage Publication.
- Taymaz, A. H. (2001). *Okulyönetimi*. Ankara: Pegem A Yayıncılık.
- Tom, A. R. (1997). *Redesigning teacher education*. New York: State University of New York Press.
- Ur, P. (1996). *A course in language teaching: Practice and theory*. Cambridge: Cambridge University Press.
- Wallace, M. J. (1991). *Training foreign language teachers: A reflective approach*. Cambridge: Cambridge University Press.

### Bio Data

**Zargham Ghapanchi** received his BA in English language and literature from Mashhad University, Iran. Then he received his MA in TEFL from Tehran Tarbiyat Modares University, Iran. He started his Ph.D. at the University of Liverpool, the UK in 1993 and Ph.D. was conferred to him in Applied Linguistics in 1998. He was the chancellor of Sabzevar Payam-e-Noor University for five years. Now he has a chair at Ferdowsi University of Mashhad. He has published several books and articles.

**Marjan Baradaran** is an M.A. student of TEFL at Ferdowsi University of Mashhad. She received her Bachelor of Science in Mathematics in 2003; however, due to her interest in English Teaching, she changed her major and started studying TEFL for master degree. She has been teaching English to adults for the past four years in Safir English Language Academy, Mashhad branch. Her research areas of interest include methodology, class management and linguistics.



## Applied ELT and Freirean Emanicipatory Education: Their Epistemological Discrepancies

Shima Ghahari<sup>1</sup>

*Ferdowsi University of Mashhad*

### Abstract

The logic behind the notion of applied ELT developed by Pishghadam (2011) is to emancipate the ELT profession from its dominated past and dependent present and to further it to a super-ordinate and independent future. Applied ELT goes beyond the typical linguistic syllabus considering life issues as against linguistic matters as its top priority. In a similar vein, Freire's (1972) emancipatory educational model espoused the idea that any kind of pedagogy should challenge the ideological, sociopolitical, and historical forces with the aim of empowering learners to function as critical agents in a society and thereby emancipating them from the hegemony of dominant ideologies (Giroux, 1988). The current study is an attempt to overview and contrast the recent idea of applied ELT and the traditional notion of critical pedagogy, suggesting that the former goes far and away ahead of the latter.

**Keywords:** Applied ELT, Emancipatory education, Critical pedagogy, Dialogical approach, Life syllabus, Education for life

### 1. Introduction

Freire's emancipatory paradigm considered education as a continuation of the hegemony of colonization in which those in power (i.e. the oppressor) aim to dominate the poor (i.e. the oppressed) for self-interest. To him, in the traditional banking model of education teachers are those who have the knowledge and are supposed to transmit it to the students who are empty vessels to be filled. The students are thereby marginalized through what he called the false generosity of the oppressor, especially in cases where teachers are from the dominant culture and students from different cultural backgrounds. The banking model then emphasizes a dichotomy between teachers and students. Freire rejects this dichotomy by his dialogical (or participatory) approach and, instead,

---

<sup>1</sup> Email: ghaharii@gmail.com



emphasizes dialogue in the tradition of Aristotle (Frankenstein & Powell, 1994; Galloway 2012; Lyons, 2001).

Although reflective language teaching largely contributes to the current practices in teacher education and is considered as the yin and yang of it, language teachers should move beyond reflective language teaching and expand their professional identities by acquiring other domains of knowledge. By so doing, learners' life style as well as language learning can be improved to a large extent. The recent theory of *applied ELT* emerged in fact with the aim of reversing the current direction, contributing to other disciplines such as linguistics, sociology, psychology, neurology rather than being contributed. Professionals in language teaching are now expected to set up training programs in which language teachers gain expertise not only in language teaching but also in other academic disciplines (Pishghadam, Zabihi, & Norouz Kermanshahi, 2012).

## 2. Theory of Emancipatory Education

Influenced by the liberation theology movement in Latin America during the 1950s and 1960s and as a response to his experience of teaching literacy with people who were marginalized due to poverty, Freire proposed the idea of *critical pedagogy*. He characterized society by relations of power and domination and, thus, called for examining issues of power, its uses, abuses, and the social relations in which it is exercised (Charaniya, Brown-Haywood, West-Welsh, & Tisdell, 2010; Mayo, 1993). The oppression that people experienced though conditions such as poverty was viewed by Freire as a fundamental process of dehumanization that turned people into objects by negating their reality and their capacity to meaningfully engage in acts of change that would give fruition to their full humanity (Morris, 2005, p. 38).

According to Freire (1972), if education is to serve the needs of the oppressed (i.e. learners), it must become a vehicle for liberation (i.e. mainstream education). Traditional education encourages a top-to-bottom approach in which the oppressed people are dehumanized and silenced by an expert (the teacher) who transmits a body of knowledge that largely represents the reality of the dominant class rather than the concerns, needs, or experiences of the oppressed. This approach "contributes to the objectification of people by rendering learners into *objects* where they engage only as passive recipients rather than as *subjects* capable of acting upon their own voice" (Morris, 2005, p. 38). The oppressed, therefore, are not encouraged to question, challenge, or see the exercise of power but, instead to accept and adapt to it, and "engage in its reproduction" (Allman, 1994, p. 6). Critical pedagogy is, therefore, grounded in "social, cultural, cognitive, economic and political context that is part of the larger community and society" (Breunig, 2005, p. 109) and encourages critical thinking and reflection in the hope for a transformative and emancipatory educational experiences (Grant &



Hurd, 2010). Emancipatory education seeks for people's wellbeing as well as the development of their potentials (Flood, 2001).

To achieve any level of emancipation brings responsibilities for both teachers and learners: Students need to engage with the materials, while teachers have a responsibility to provide instruction in an appropriate format. This kind of interaction between student and teacher can certainly contribute to the quality of teaching and learning outcomes. Teachers, however, have additional responsibilities as they are in a position of power. They must be aware of the influence their beliefs and attitudes may have on others (Grant & Hurd, 2010). For example, according to Moon, reflective practice of teachers facilitates the processes of learning, and assists them to not only continue their professional development, but also model cycle of experiential learning to students (1999). Education, therefore, as a form of power and as a medium for emancipation, can be used or misused to influence liberatory outcomes; it can bring about physical, mental, emotional, and spiritual consequences (Kember & Kwan, 2002). Those who aspire to *teach* what others may *learn* must be cognizant of the responsibilities involved, and the potential for unfair distributions of power (Perriton & Reynolds, 2004). In order to facilitate learning, teachers must help students develop skills and processes relevant to their course of study as well as everyday life. Students must be encouraged to look beyond common assumptions and stereotypes and instead try to apply a critical perspective in a constructive manner. Such an attitude has led teachers to a pedagogy of co-enquiry with students, and a continual reflection on teaching practices. If teachers, thus, intend to help the oppressed (i.e. learners) in finding true ways of acting on the conditions of oppression, Allen (2002) proposes, they must believe, "situationally and strategically, in the ontology, epistemology, and axiology of the oppressed" (p. 23).

Freire's method of emancipatory education thus incorporates three principal components: a) the reality of the oppressed (i.e. learner), b) the process of conscientization that occurs through dialogue and group analysis of the conditions that give rise to oppression, and c) praxis (Freire, 1972). The model utilizes participatory learning methods and a problem-posing approach; learners are encouraged to deepen their understandings of reality through critical reflection on the personal and sociopolitical dimensions of the problems they are encountering.

Problem-posing is a three-stage methodology that involves listening to understand the felt issues or themes of concern in the group, collectively engaging in participatory problem-solving dialogue around a *code* that physically represents the identified group issue through strategies such as role plays, stories, and photographs, and engaging in action (Morris, 2005, p. 40).



Critical reflection does not take place in a vacuum, however. The educator needs to adopt a sequential questioning strategy and moves discussion from general, through to personal, social analysis, and action. The sequencing is important in that it represents the extent to which the learners' values and beliefs are challenged and critical reflection deepened. The outcome of critical reflection is conscientization: As the oppressed (i.e. learners) become cognizant, they recognize the links between structural forces and oppression. Meanwhile, the learners experience empowerment as they shift from being passive *objects* that are acted upon to being *subjects* responsible for their own destiny. Finally, the outcome that any transformational learning might bring about is engagement in emancipatory praxis to shape the contextual reality of people's lives (Burbules & Berk, 1999; Morris, 2005).

Overall, pedagogy in the critical sense illuminates the relationship among knowledge, authority, and power (Giroux, 1994). This epistemology is in direct opposition to the positivist paradigm in educational theory. "Positivists view knowledge as neutral, value-free, and objective, existing totally out-side human consciousness"; students are passive recipients of that knowledge and are to discover these static facts and classify them. Freire's dialogical approach, however, seeks for the interaction between teacher and learner in a way that can contribute to the process of knowledge building. Such a process, according to Freire, is unlike the banking education model which involves transmission of ready-made knowledge from teacher to student through the activity of teaching. Freire insists that "knowledge is not static; that there is no dichotomy between objectivity and subjectivity, or between reflection and action; and that knowledge is not neutral" (Frankenstein, 1983, p. 316). To him, teachers raise problems and then both groups (i.e. students and teachers) discuss them in order to bring valuable opinions and abilities to the learning. Teachers do not provide answers, but work with students to develop *critical thinking* or *conscientization*. They work together to investigate a different reality. By so doing, the oppressed are entitled to consider their reality as a problem to be transformed.

### 3. Theory of Applied ELT

According to Pishghadam (2011), English language teaching (ELT), just like psychology, neurology, mathematics, and computer sciences, can be considered as having two aspects of theoretical and applied. While underscoring the theoretical aspect of ELT, ELT theorizers and practitioners have disregarded its applied part. Traditionally, ELT has been considered as a sub-branch of applied linguistics, which requires its prescriptions and proscriptions to improve its own status. On the other hand, ELT practitioners and teachers have been merely consumers of the findings of disciplines like linguistics. Today, however, the trend has changed: ELT theorizers and practitioners are no more consumers but autonomous. "Now it is time for them to play a producer role" and for the applied



ELT to take "a more contributory role". According to Pishghadam, "ELT has grown in maturity over years, establishing an independent identity for itself. It does not play second fiddle to applied linguistics anymore" (2011, pp. 9-11).

Applied ELT has been operationalized in some sample studies. Pishghadam (2008), for instance, has suggested that literary discussion in ELT classes can enhance the critical thinking abilities of the learners (Pishghadam, 2011; Pishghadam & Naji, 2012). Similarly, Khazaifar, Pishghadam, and Ziai (in press) indicated that English language reading materials could be designed to improve critical abilities. The findings of Hosseini, Pishghadam, and Navari (2010) have revealed that a language learning class can increase emotional intelligence competencies of learners via managing their anxiety and fostering interpersonal competencies (Pishghadam, 2011). In another study, Pishghadam and Saboori (2011) showed that those English language teachers in Iran who held positive attitudes towards the American culture and used to act native-like tended to alienate students from their own home culture. The study, in fact, supported the pivotal role language teachers play in fostering national identity of the learners. "Therefore, English language learning classes have the potential to be the sites for developing the cultural and national identity of the learners" (Pishghadam, 2011, p. 11).

Closely related to Pishghadam's notion of applied ELT is his new type of syllabus called *life syllabus*. Whereas the traditional syllabi focus on different linguistic features to be touched on in ELT classes, life syllabus gives the top priority to life issues rather than language. It in fact allows language teachers to take advantage of the full potentials as well as unique features of ELT classes to first reinforce learners' motivation, critical thinking, creativity, among other life qualities, and then to teach a language. By so doing, teachers help learners overcome psychological barriers such as stress, demotivation, depression, burnout, and uncertainty avoidance in advance of learning a new language (Pishghadam, Zabihi, & Norouz Kermanshahi, 2012).

Another significant aspect of applied ELT is that it goes beyond the typical linguistic syllabus considering life issues as against linguistic matters as its first priority. This new syllabus is dubbed as *life syllabus*. Current syllabus design involves selecting and sequencing linguistic features varying from grammar and vocabulary (i.e. structural syllabus), notions and functions (i.e. notional-functional syllabus), and tasks (i.e. procedural syllabus and task-based method). These approaches, however, have numerously questioned for failing to address learners' actual communicative needs and to wrongly view second language learning as a linear process (Baleghizadeh, 2008). In this new syllabus, however, language learning is more educationally oriented and revolves more around the most significant life issues (Pishghadam, 2011). A number of educational philosophers like Dewey (1897), Freire (1998), Krishnamurti (1981), and Walters



(1997) have also emphasized on the importance of life issues in education and stipulated that any educational system must address and meet its educators' life challenges, improve their some, not to say all, aspects of their lives, and thereby prepare them for lifelong learning process (Pishghadam & Zabihi, 2012). For Noddings (2003), for instance, individual's happiness must be the aim of education; Walker (1999) considers self-determination as the primary goal; Matthews (2006) emphasizes on the improvement of emotional abilities; and still others like Hare (1999) and Winch (1999) prioritize critical thinking and individual's autonomy, respectively. ELT is no exception. Given the vast number of its potentials as well as the primary deficiencies of the current ELT syllabi which, at their very best, take a single (or more) aspect of language as their desired outcome to be taught and mastered in a cumulative fashion, life syllabus which directs is a more purposeful and rewarding alternative.

Moreover, ELT classes have a very different atmosphere in which lots of human abilities can be nurtured in addition to language learning. According to Pishghadam (2011),

- a) In these classes a number of topics are discussed; these topics include all things about life, music, culture, politics, society, science, etc. In addition to informative role of the topics, they can provide the learners with more food for thought about life. Based on these discussions, learners can reflect more on their life, trying to change the status quo.
- b) They are characterized by the overuse of pair works and group works. This way of interaction and exchange of information can help the dynamicity of the class, enhancing different communicative abilities in students. The situation makes the class be more dialogic and free from more monologic and narrative classes.
- c) In EFL/ESL classes the learners face another culture more directly. By reading another culture, they get more acquainted with their own home culture. This situation can help the schizophrenic nature of the class, in which two big personalities (cultures) struggle for more dominance. Home culture and the foreign culture can have dialogue in class, helping the learners to create or maintain identities.
- d) In EFL/ESL classes students get more familiar with the structures and words of another language. According to Vygotsky, grammar is a tool for enhancing the higher order abilities of human beings. Grammar can make individuals think more logically. Besides, words of two languages are different which may convey different types of understanding, shedding more light on our vague grasp of things. Thus this linguistic contact of two languages makes the setting more unique.
- e) Another important feature of English language learning classes is that students might take more freedom to express themselves and show their real self. By speaking in another language, one can project his/her own true identity, especially in nondemocratic settings.



When one speaks in another language, he/she feels free to say something they cannot express in their mother tongue owing to social or political reasons.

f) Last is the fun nature of EFL/ESL classes. Watching movies, listening to different songs, discussing different topics, using computers, the Internet, mobiles, and different kinds of tasks make the class be a fun. Fun and learning come together to create an embracing atmosphere in class.

These unique features characterize ELT classes as settings wherein a life syllabus, that is teaching language along with other life issues, should be adopted and one must first develop "critical abilities, creativity, social intelligence, emotional intelligence, and then [learn] a language" (Pishghadam, 2011, p. 13).

#### **4. Discussion**

Since the postmodern movement, unlike the pre-modern and modern eras, teachers are no more regarded as passive executors to carry out the prescriptive rules of language teaching, but are encouraged to reflect upon their teaching on the basis of the context in which it occurs. (For a full review of modernism and postmodernism see Pishghadam & Mirzaee, 2008). Language teachers are now reflective practitioners who need to evaluate their teaching through action research and adapt their teaching practices to the requirements of the contexts of teaching and learning at hand (Pishghadam, Zabihi, & Norouz Kermanshahi, 2012). Meanwhile inspired by the ideas of Freire (1972), scholars like Kumaravadivelu (1999) and Pennycook (1999), among others, sought for a normative educational project which addresses social, economic, and cultural inequality. They cried out for analyzing, critiquing, and altering inequitable knowledge structures and social relations of school and society: "language teaching and learning is an act of political and cultural power with substantive material and social consequences and possibilities for learners and their communities" (Luke & Dooley, 2011, p. 856). In fact, under the rubric of critical pedagogy, language teachers were assigned a critical and participatory role and turned to *transformative intellectuals*

whose primary obligation was to empower language learners and emancipating them from the hegemony of dominant ideologies...[Language teachers are then viewed as] agents of change and seekers of democracy in education, should therefore raise their sociopolitical awareness via problem-posing activities in terms of a holistic approach to both educational advancement and personal transformation in such a way that teachers and learners acquire a sense of ownership of their own teaching and learning rather than unduly relying on professional experts (Pishghadam, et al., 2012, pp. 894-895).



Now, inspired by but not limited to Freire's emancipatory and dialogical forms of education, applied ELT has emerged as a new paradigm in second language learning and teaching. It aims to entitle ELT an autonomous status to contribute to rather than be contributed by other disciplines such as psychology and sociology, giving it a more life-changing status (Pishghadam, 2011). In this new paradigm,

- 1) Language teachers need to adopt a new role; they are expected to become *educational language teachers*, that is, experts who try "to improve other domains of knowledge which, directly or indirectly, affect learners' idiosyncratic lives" (Pishghadam et al., 2012, p. 895). Language teachers are required, therefore, to expand their knowledge of other disciplines so as to understand the psychological, emotional, social, economic, religious, and moral needs of learners.
- 2) The focus is shifted away from a language syllabus (i.e. *language-and-life* classes) towards life syllabus (i.e. *life-and-language* classes) wherein language is at the service of enhancing life qualities. Language teachers, therefore, are required to give the top priority to learners' life issues and then to teach the desired language.
- 3) Syllabus designers and material developers can think of ways to design and implement *life syllabuses*, by diagnosing the areas that need to be targeted. In fact, ELT classes can be places wherein some qualities of life are enhanced and promoted and also language needs of learners are tailored to their specific contexts. By so doing, ELT can be a constructive discipline to help periphery countries (i.e. the oppressed) reduce dependency on center countries (i.e. the oppressor) and have a voice in the age of globalization where English language plays a significant role (Pishghadam & Naji, 2012).
- 4) *English for life purposes* (ELP) has been proposed to provide the learners with the chance to compare their home culture and identity with other cultures and to develop their unique identities. It missions to alleviate the learners' anxiety, depression, demotivation, shyness, or other negative aspects of life and, instead, improve their emotional, intellectual, spiritual, and motivational abilities through teaching them a second/foreign language.
- 5) *World Englishes* and *multi-competence* are emphasized: Avoiding learners' mother tongues, presentation of target culture, use of authentic materials, and acquiring a native-like accent are no more goals for applied ELT to pursuit. In fact, textbooks are culturally rather than linguistically authenticated (Pishghadam & Naji, 2012).



- 6) Learners' characteristics such as motivation, emotional abilities, cognitive and affective styles and values are no more considered as obstacles for language teaching and learning profession, but as real strength points and a set of what Pishghadam (2011) regards as the unique features of language classes.
- 7) Communicative language teaching (CLT) can be appropriated to the local contexts. CLT has been attacked for being an import of linguistic imperialism to the extent that Phillipson labeled it as *communicative imperialism* (2009): It incorporates tasks and methods developed by western countries with an attempt to impose their cultures as well as thinking and living systems on the underdeveloped societies or what Freire calls *the oppressed*. Applied ELT, however, provides the ground for localizing ELT material in favor of one's native culture thereby thwarting the cultural imperialism brought about by current trends in ELT. Moreover, the unique features of language classes discussed above, once localized to specific contexts and cultures, can help enrich and further learners' native cultures (Pishghadam & Naji, 2012).
- 8) Unlike current trends of ELT, national and cultural identities of learners are developed and upheld. Teachers are trained to deal with familiar issues and learners' home culture and, thereby, to let them invest something personal in the learning process. According to Cummins (1994), inclusion in language course books of local cultures as well as a familiar content, as is the main objective of applied ELT, facilitates the learning process and empowers learners' identity through the positive contribution of Zone of Proximal Development (ZPD). In so doing, language learning is no more a means of linguistic imperialism or an evidence for neocolonialism, but as a gateway for learners' cultural enrichment while keeping pace with the globalized world.

## References

- Allen, R. L. (2002). Wake up, Neo: White consciousness, hegemony, and identity in *The Matrix*. In: J. Slater, S. Fein & C. Rossatto (Eds.), *The Freirean legacy: Educating for social justice* (pp. 104-125). New York: Peter Lang Publishers.
- Allman, P. (1994). Paulo Freire's contributions to radical adult education. *Studies in the Education of Adults*, 26, 144-161.
- Baleghizadeh, S. (2008). Task-supported structural syllabus. *Journal of Curriculum Studies*, 3, 8-22.
- Breunig, M. (2005). Turning experiential education and critical pedagogy theory into praxis. *Journal of Experiential Education*, 28(2), 106-122.
- Burbules, N. C., & Berk, R. (1999). Critical thinking and critical pedagogy: Relations, differences, and limits. In Thomas S. Popkewitz and Lynn Fendler (Eds.) *Critical theories in education*. Routledge: New York. Available at <http://faculty.ed.uiuc.edu/burbules/ncb/papers/critical.html>



- Charaniya, N., Brown-Haywood, F., West-Welsh, J., & Tisdell, E. J., (2010). *The intersecting roles of religion, culture and spirituality in emancipatory adult education*. Theory Development proposal presented at the Adult Education Research Conference.
- Cummins, J. (1994). Knowledge, power, and identity in teaching English as a second language. In F. Genesee (Ed.), *Educating second language children* (pp. 33-58). Cambridge: Cambridge University Press.
- Dewey, J. (1897). My pedagogic creed. *The School Journal*, 54(3), 77-80.
- Flood, R. L. (2001). The relationship of 'systems thinking' to action research. In P. Reason & H. Bradbury (Eds.), *Handbook of action research*. London: Sage Publications.
- Frankenstein, M. (1983). Application of Paulo Freire's epistemology. *Journal of Education*, 165, 315-339.
- Frankenstein, M. & Powell, A. (1994). Toward liberatory mathematics: Paulo Freire's epistemology and ethno-mathematics. In Mc Laren, P. L. & C. Lankshear (Eds.), *Politics of liberation: Paths from Freire* (pp. 74-99). London: Rutledge.
- Freire, P. (1972). *Pedagogy of the oppressed*. Harmondsworth: Penguin.
- Freire, P. (1998). *Teachers as cultural workers - Letters to those who dare teach*. Westview Press, Boulder, CO.
- Galloway, S. (2012). Reconsidering emancipatory education: Staging a conversation between Paulo Freire and Jacques Rancière. *Educational Theory*, 62, 163-184.
- Giroux, H. A. (1988). *Teachers as intellectuals: Toward a critical pedagogy of learning*. South Hadley, Massachusetts: Bergin & Garvey.
- Giroux, H. A. (1994). *Disturbing pleasures: Learning popular culture*. New York: Routledge.
- Grant, S. & Hurd, F. (2010). Incorporating critical pedagogy into the scholarship of teaching and learning: Making the journey alongside our students. *International Journal of the Scholarship of Teaching and Learning*, 4, 1-12.
- Hare, W. (1999). Critical thinking as an aim of education. In R. Marples (Ed.), *The aims of education* (pp. 85-99). London: Routledge.
- Hosseini, A., Pishghadam, R., & Navari, S. (2010). Tasire classhaye zaban dar afzayesh hooche hayajani. *Language and Literature Studies*, 42, 1-11.
- Kember, D., & Kwan, K. P. (2002). Lecturer's approaches to teaching and their relationship to conceptions of good teaching. In N. Hativa & P. Goodyear (Eds.), *Teacher thinking, beliefs and knowledge in higher education* (pp. 219-239). Dordrecht, the Netherlands: Kluwer Academic Publishers.
- Khazaifar, A., Pishghadam, R., & Ziai, S. (in press). Ahamiyate tafakore enteghadi dar khandan. *Farhangestan Letter*.
- Krishnamurti, J. (1981). *Letters to the schools*. Madras, India, KFI.



- Kumaravadivelu, B. (1999). Theorizing practice, practicing theory: Critical classroom observation. In H. Trappes-Lomax and I. McGrath (Eds.) *Theory in language teacher education* (pp. 33-45). London: Prentice-Hall.
- Luke, A., & Dooley, K. (2011). Critical literacy and second language learning. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (pp. 856-868). London: Routledge.
- Lyons, J. (2001). *Paulo Freire's educational theory*. Retrieved from [www.newfoundations.com/gallery/Freire.html](http://www.newfoundations.com/gallery/Freire.html)
- Matthews, B. (2006). *Engaging education: Developing emotional literacy, equity, and co-education*. Buckingham, UK: Open University Press.
- Mayo, P. (1993). When does it work? Freire's pedagogy in context. *Studies in the Education of Adults*, 25, 11-30.
- Moon, J. (1999). *Reflection in learning and professional development. Theory and practice*. London: Kogan Page Ltd.
- Morris, M.A. (2005). *Utilizing the theories of mezirow and freire to foster learning of the principles of health promotion during international nursing practice experiences*. Unpublished MA thesis at the University of British Columbia.
- Noddings, N. (2003). *Happiness and education*. Cambridge: Cambridge University Press.
- Pennycook, A. (1999). Introduction: Critical approaches to TESOL. *TESOL Quarterly*, 33, 329-348.
- Perriton, L., & Reynolds, M. (2004). Critical management education: from pedagogy of possibility to pedagogy of refusal? *Management Learning*, 35(1), 61.
- Pishghadam, R. (2008). Afzayeshe tafakore enteghadi az tarighe mobahesye adabi. *Journal of Literature*, 48, 153-167.
- Pishghadam, R. (2011). Introducing applied ELT as a new approach in second /foreign language studies. *Iranian EFL Journal*, 7, 8-14.
- Pishghadam, R., & Mirzaee, A. (2008). English language teaching in postmodern era. *TELL*, 2, 89-109.
- Pishghadam, R. & Naji, E. (2012). Applied ELT as a panacea for linguistic imperialism. *Iranian EFL Journal*, 8, 35-52.
- Pishghadam, R., & Saboori, F. (2011). A Qualitative analysis of ELT in the language institutes of Iran in the light of the theory of 'World Englishes'. *Journal of Language Teaching and Research*, 2, 569-579.
- Pishghadam, R. & Zabihi, R. (2012). Life syllabus: A new research agenda in English language teaching. *TESOL Arabia Perspectives*, 19, 23-27.
- Pishghadam, R., Zabihi, R., & Norouz Kermanshahi, P. (2012). Educational language teaching: A new movement beyond reflective/critical teaching. *Life Science Journal*, 9, 892-899.



- Walker, J. C. (1999). Self-determination as an educational aim. In R. Marples (Ed.), *The aims of education* (pp. 112-123). London: Routledge.
- Walters, J. D. (1997). *Education for life: Preparing children to meet the challenges*. US, Nevada: Crystal Clarity Publishers.
- Winch, C. (1999). Autonomy as an educational aim. In R. Marples (Ed.), *The aims of education* (pp. 74-84). London: Routledge.

### Bio Data

**Shima Ghahari** is a Ph.D. candidate of TEFL at Ferdowsi University of Mashhad and is currently teaching at the same university. She received her M.A. in TEFL from University of Tehran and studied English Language Translation for her B.A. at Allameh Tabatabai University. She has contributed many publications in different domains of English language education to journals and conference proceedings. Her major research interests include Sociology and Psychology of language education as well as bilingualism and cultural studies.



## Investigating Academic Discipline Bias in UTEPT Using MH Method

Seyyed Mohammad Reza Amirian<sup>1</sup>

*Hakim Sabzevari University, Iran*

### Abstract

The purpose of the present bias study was to detect the items performing differentially in the University of Tehran English Proficiency Test (UTEPT) using Mantel-Haenszel (MH) method. The participants of the study consisted of 1550 test takers with humanities and science and engineering academic backgrounds. MH method, as one of the most popular DIF detection methods, was employed to examine whether test taker's academic discipline resulted in their differential performance. After identifying DIF items, a content analysis of the items with substantial effect size was conducted in order to uncover the underlying sources of the differential performance of DIF items. The findings of the study revealed that, overall, 13 items exhibited DIF. Moreover, the content analysis of DIF items with sizable magnitude demonstrated that only 4 items displayed academic discipline bias. As a result, it is concluded that UTEPT is fair to both humanities and science and engineering.

**Keywords:** Item bias, DIF, DIF effect size, UTEPT, MH

### 1. Introduction

The issue of test bias has attracted considerable attention in recent years. Item bias is said to occur when examinees of one group are less likely to answer an item correctly than examinees of another group because of some characteristic of the test item or testing situation that is not relevant to the test purpose.

Item bias has considerable ramifications at a policy, administrative, and classroom level because bias can lead to systematic errors that distort the inferences made in classification and selection of students (Zumbo, 1999). Learners who have similar knowledge of the material on a test, based on total examination results, should perform similarly on individual examination items, regardless of gender, culture, ethnicity, or race.

---

<sup>1</sup> Email: smr.amirian@gmail.com



The first logical step in detecting bias is to find items where one group performs much better than the other group. Such items function differentially for the two groups and this is known as *Differential Item Functioning (DIF)*. DIF identifies test items that function differentially for two groups of test takers, without the discriminatory overtones of the term bias.

Zumbo (1999, p. 12) holds that "DIF occurs when examinees from different groups show differing probabilities of success on (or endorsing) the item after matching on the underlying ability that the item is intended to measure". In other words, DIF is said to be present when the probabilities of success on a given item are variant between the two groups at the same ability level (Kim, 2001). A DIF item may be considered biased when a score difference between two or more groups is due to a factor that is not the construct being tested.

It is important to note that DIF is a necessary but not sufficient condition for bias. Bias only exists if the difference is illegitimate, that is, if both groups should be performing equally well on the item. An item may show DIF but not be biased if the difference is due to actual differences in the groups' ability needed to answer the item which is called *impact*. For example, if one group is high proficiency and the other low proficiency: The low proficiency group would necessarily score much lower. Only where the difference is caused by construct-irrelevant factors can DIF be viewed as bias.

In other words, an item that shows DIF needs to be investigated further to uncover the reasons for its differential functioning. Most DIF analyses compute DIF for a potentially disadvantaged group (known as the focal group) compared to the potentially advantaged group (known as the reference group).

Once identified, DIF may be attributed either to item bias or to item impact. "Item bias is defined as invalidity or systematic error in how a test item measures a construct for the members of a particular group" (Camilli & Shepard, 1994, p. 8). It is systematic because it constantly distorts performance for members of the group. Alternatively, group disparity in item performance that reflects actual knowledge and experience differences on the construct of interest is item impact.

DIF is of uniform and non-uniform nature. Uniform DIF occurs when there is no interaction between the ability level and group membership. Non-uniform DIF exists when the probability of correctly answering an item is higher for one group at some points on the scale, and higher for the other group at other points. In other words, if there is an interaction between ability and group, this signals the existence of non-uniform DIF (Swaminathan & Rogers, 1990).



### *1.1. Mantel-Haenszel*

Mantel-Haenszel (MH) as a practical technique for determining if a test item is functioning differentially for two groups of examinees was first proposed by Holland (1985). Later, Holland and Thayer (1988) introduced MH as a DIF technique. MH is arguably the most widely used contingency table approach to studying DIF (Clauser & Mazor, 1998).

MH is a nonparametric approach to identifying DIF which explicitly matches the examinees from two different groups on the ability of interest, and then compares the likelihood of success on the item for the two groups across the score scale. Typically, total test score is used as the matching criteria to group examinees. The null hypothesis for the MH statistic states that the odds for the focal group answering the test item correctly is the same as the odds for the reference group. Together with standardization, MH is the primary DIF detection procedure used in Educational Testing Service (Liu, Harris, & Schmidt, 2007).

Rogers and Swaminathan (1993) consider the MH procedure as one of the most popular procedures for detecting DIF. They believe that the primary reasons for its popularity include its computational simplicity, ease of implementation, and associated test of significance. However, MH is designed to detect uniform DIF and may not be appropriate for detecting non-uniform DIF. The MH procedures belong to contingency tables procedures, together with Logistic Regression (LR), log linear models and some simpler indices. The MH test statistic tests the null hypothesis of no relation between group membership and test performance on an item after controlling for ability (usually in terms of overall test performance). The test is based on the odds ratio between correct and incorrect responses, between a reference and a focal group when conditioning on total test score.

## **2. Method**

### *2.1. Participants*

The data for the present study were gathered from 1550 test takers who took UTEPT in 2010. The sample was divided into a reference group of 809 test takers with humanities background as reference group and 741 test takers with science and engineering background as focal group. The humanities group consisted of applicants of social sciences, law, political sciences, management, Persian literature, and foreign languages, and the science and engineering group was comprised of students of chemistry, physics, mathematics, biology, agricultural engineering, mechanical engineering, electrical engineering, and civil engineering. They were all Ph.D. candidates seeking to



sit for Ph.D. exams of the University of Tehran. The participants of the study were both male and female and of various age groups.

## *2.2. Instruments*

### *2.2.1. UTEPT*

University of Tehran English Proficiency Test (UTEPT) is a high stake test of English developed and administered by University of Tehran (UT). Passing UTEPT is a prerequisite for master's degree holders aiming at participating in PhD exams of UT. UTEPT is administered four times a year with about 2000 participants taking it each time. It consists of 100 multiple choice items with a time limit of 100 minutes. The test is composed of three sections: 1) Structure and written expression including 15 structure items, 10 written expression items and 5 grammar-in-context items, 2) vocabulary consisting of 30 synonym items and 5 vocabulary-in-context items, and 3) reading comprehension encompassing 26 reading comprehension and 4 restatement items. The reading section covers 6 passages each followed by 4 to 8 questions. There are both short and long passages with a range of 95 to 359 words.

### *2.2.2. DIF questionnaire*

A DIF questionnaire, adapted from Geranpayeh and Kunnan (2007), was used in the content analysis phase of the study. The questionnaire was supposed to elicit judgments from two expert judges on whether UTEPT advantaged a particular academic discipline group. The questionnaire consisted of a 5-point scale from 1 (strongly advantage) to 2 (advantage) to 3 (neither advantage nor disadvantage) to 4 (disadvantage) to 5 (strongly disadvantage). The experts were asked to give their comments as to why they believed an item advantaged or disadvantaged a particular academic discipline group.

### *2.2.3. DIFAS program*

DIFAS (Penfield, 2009) performs a variety of functions related to assessing the presence of DIF. The DIF procedures that DIFAS runs for dichotomously scored items include:

#### *Mantel-Haenszel Chi-Square (MH CHI) –*

The Mantel-Haenszel chi-square statistic (Holland & Thayer, 1988) is distributed as chi-square with one degree of freedom.



*Mantel-Haenszel Common Log-Odds Ratio (MH LOR) –*

The Mantel-Haenszel common log-odds ratio (Camilli & Shepard, 1994) is asymptotically normally distributed. Positive values indicate DIF in favor of the reference group, and negative values indicate DIF in favor of the focal groups.

*Breslow-Day Chi-Square (BD) –*

The Breslow-Day chi-square test of trend in odds ratio heterogeneity (Breslow & Day, 1980; Penfield, 2003) is distributed as chi-square with one degree of freedom. This statistic has been shown to be effective at detecting non-uniform DIF.

*Combined Decision Rule (CDR) –*

The combined decision rule (CDR) flags any item for which either the Mantel-Haenszel chi-square or the Breslow-Day chi-square statistic is significant at a Type I error rate of 0.025 (Penfield, 2003). The message OK is printed if neither statistic is significant, and the message FLAG is printed if either statistic is significant.

*The ETS Categorization Scheme (ETS) –*

The ETS categorization scheme (Zieky, 1993) categorizes items as having small (A), moderate (B), and large (C) levels of DIF.

### 3. Results

The findings of the study indicated that out of 100 items, 13 items (13%) displayed academic discipline DIF. Items flagged with DIF include items 10, 33, 35, 38, 39, 40, 47, 50, 52, 55, 69, 73, and 78. Concerning the specific sections of these items, it is realized that items 10, 33, and 35 come from the grammar section, items 38, 39, 40, 47, 50, 52, and 59 come from the vocabulary section, and items 69, 73, and 78 come from the reading comprehension section (Table 1).

**Table 1.**

MH academic discipline DIF findings in different sections of UTEPT

Section	Number	Percent
Grammar	3	3%
Vocabulary	7	7%
Reading Comprehension	3	3%
Total	13	13%



As indicated in Table 1, the vocabulary section of UTEPT was found to contain seven DIF items which is more than 50 percent of all academic DIF items. This means that the vocabulary section of UTEPT is mainly responsible for differential performance of this test which is somehow contrary to our expectations since normally it is the reading comprehension section of language test that leads to their differential performance. It is the reading passages that contain texts that are long enough to cover a subject which is advantaging a particular academic discipline.

Considering the magnitude of DIF items, it is found that only four items manifest moderate effect size according to ETS classification while nine items show negligible DIF magnitude (Table 2). This shows that out of 100 items on UTEPT, there are only 4 items which display substantial magnitude of DIF while all other items are functioning well and do not substantially advantage one group of the participants over the other.

**Table 2.**

MH academic discipline DIF results

Item	No.	MH CHI	MH LOR	BD	CDR	ETS
Grammar	10	6.0846	-0.3401	1.009	Flag	A
Grammar	33	19.8096	-0.4959	0.008	Flag	B
Grammar	35	1.8819	-0.1595	8.573	Flag	A
Vocabulary	38	6.1336	0.3703	3.376	Flag	A
Vocabulary	39	10.1713	-0.4213	0.06	Flag	A
Vocabulary	40	16.8032	0.6211	0.006	Flag	B
Vocabulary	47	6.1015	0.2791	17.17	Flag	A
Vocabulary	50	20.2237	0.59	0.043	Flag	B
Vocabulary	52	27.6599	0.6763	1.339	Flag	B
Vocabulary	55	12.2867	0.3895	0.073	Flag	A
Reading	69	0.6817	0.1213	10.863	Flag	A
Reading	73	9.2183	-0.365	0.148	Flag	A
Reading	78	9.0135	-0.3854	1.514	Flag	A

As Table 2 shows, the moderate size DIF items are items 33, 40, 50, and 52. Item 33 is a grammar item and items 40, 50, and 52 are vocabulary items. No item in reading comprehension section displays considerable DIF effect size. Category-A DIF items or items with negligible effect size include items 10, 35, 38, 39, 47, 55, 69, 73, 78. In terms of their relevant sections, it is shown that



items 10 and 35 belong to the structure section, items 38, 39, 47, and 55 belong to the vocabulary section and items 69, 73, and 78 belong to the reading comprehension section.

As far as the direction of DIF is concerned, based on the index of MH LOR, it is shown that seven items containing items 38, 40, 47, 50, 52, 55, and 69 favor the reference group (humanities) while six items including items 10, 33, 35, 39, 73, and 78 are in favor of the focal group (science and engineering). Only item 33 with moderate effect size favors science and engineering while three items, that is, 40, 50, and 52, with moderate effect size favor the humanities group. Item 33 is a grammar in context item, items 40 and 50 are vocabulary in sentences, and item 52 is a fill in the blank vocabulary item.

### *3.1. Content analysis*

Once DIF items are detected, it is crucial to investigate the sources of DIF across the groups; because, DIF items are not necessarily biased items. As Clauser and Mazor (1998) point out, DIF is a necessary but not a sufficient condition for demonstrating item bias. Results of the DIF analysis should lead to the investigations of the potential sources of bias. The sources of DIF may be examined in terms of two aspects: the test characteristics and the test-takers' characteristics. If the source of DIF is found in test characteristics, the test can be said to be biased for a particular group, as indicated by the DIF results (Kim, 2001).

According to MH findings, four academic discipline DIF items revealed moderate DIF magnitudes based on ETS classification scheme including items 33, 40, 50, and 52. In order to examine the sources of DIF with these items, a DIF questionnaire developed by Geranpayeh and Kunnan (2007) was adapted. Accordingly, two content experts were asked to rate the suitability of the test items for academic discipline group using a questionnaire with a 5-point scale. The researcher briefed the experts about the nature of the task and what was expected of them.

Based on content experts' judgment, out of four academic discipline DIF items with sizable DIF effect sizes, three items appeared to advantage humanities and one item appeared to advantage science and engineering. Item 33 is the only item that showed DIF to the advantage of science and engineering and items 40, 50, and 52 displayed DIF to the advantage of humanities. Below, the contents of these items are analyzed to investigate the potential sources of DIF.

**Item 33.** *The emergence of endocrinology as a separate discipline can be .....32.....This .....33.....is secreted from cells in the intestinal walls when food.....34..... the stomach.*



A. *substance*

B. *spray*

C. *solution*

D. *subject*

Item 33 is a grammar in context (cloze test) question where students are supposed to fill in the blank with one of the four choices suggested for the gap. Examinees are supposed to choose A, “*substance*”, as the word that completes the sentence. This item comes from a cloze test which is on human biology, a scientific text. The reason why this item shows DIF against humanities might be the fact that the word “*substance*” is more likely to be encountered in scientific texts and examinees with science and engineering background are more probable to have seen this word in their textbooks. This is confirmed by the results of content analysis done by content specialists. Both content experts believed that this item advantages science and engineering group because the passage is on a scientific topic. Consequently, the source of considerable DIF magnitude for this item can be attributed to academic discipline bias.

**Item 40.** *There is evident conflict between Henry's social philosophy and the actions of his characters.*

A. *general*

B. *obvious*

C. *important*

D. *appropriate*

This vocabulary-in-sentence item aims at measuring test takers' knowledge of synonyms. Choice B, “*obvious*”, must be marked as the best synonym for the adjective “*evident*”. MH flagged this item as being differentially easier for humanities. Although this structure item gives a very short context, the fact that it talks about social philosophy might be hypothesized as a possible source of bias. This argument is supported by content experts who contended that this item favors humanities. Probably examinees in humanities group are advantaged because they have a higher chance of seeing the phrase “*evident conflict*” in their textbooks. Thus, the potential source of DIF for this item can be related to academic discipline bias rather than their language ability that is impact.

**Item 50.** *As paradoxical as it may seem, the infinity of even numbers is as big as that of all numbers.*



- A. mathematical
- B. contradictory
- C. methodological
- D. rudimentary

Like the previous item, this vocabulary item intends to assess the knowledge of synonyms in English. The word “*contradictory*” must be chosen as the best synonym for “*paradoxical*”. It was found that examinees coming from humanities background performed significantly better on this item. In other words, this item systematically advantages humanities over science and engineering. One of the content experts argued that this item advantages humanities because the word “*paradoxical*” is more likely to appear in humanities text whereas the second expert believed that this item neither advantaged humanities nor science and engineering. As the words “*paradoxical*” and “*contradictory*” are possibly more frequent in humanities fields, it can be hypothesized that the source of DIF for item 50 is in the content of test. That is to say item 50 is probably biased against science and engineering group.

**Item 52.** *Many people in the past could neither read nor write. They were.....*

- A. illiterate
- B. traditional
- C. emotional
- D. cultural

For this item, the test takers were supposed to mark “*illiterate*” as the best item that matches the definition. Given that this item works to the advantage of humanities, it seems reasonable to conjecture that since the word “*illiterate*” is more likely to be encountered in humanities texts, it is in favor of humanities. Thus, it can be hypothesized that this item functions differentially against science and engineering group due to their discipline and not their language ability. That is to say, this item is biased against science and engineering group.

#### 4. Discussion

The findings of the study revealed that out of 100 UTEPT items, MH flagged 13 items (i.e. 13%) as performing differentially for humanities and science and engineering groups. This shows that academic discipline DIF is prevalent on UTEPT items, which assumes importance considering the decisions we intend to make about the academic future of examinees based on the information



provided by UTEPT. This finding aligns with Hale's (1988) findings who found that major-field area affected the reading performance of TOEFL test takers, although the effects were not large.

Considering the direction of DIF items, it was revealed that out of a total of 13 MH uniform DIF items, seven items were differentially easier for humanities whereas six items were easier for science and engineering. This finding indicated that relatively similar number of items was in favor of humanities and science and engineering, as a result the differential effect of discipline is cancelled out throughout the whole test.

Therefore, the individual direction of DIF indices and the cumulative effect of these indices altogether displayed random preferential treatment for either of the academic groups. As a result, the designers and users of UTEPT are reassured that, overall, the test does not systematically favor one academic discipline group over the other.

Let us not forget, however, that the test of significance alone cannot be an indicator of significant DIF and accompanying a test of significance with a measure of effect size can shed more light on the magnitude of DIF values.

The findings indicate that the vocabulary section of UTEPT incorporates three substantial academic discipline DIF items. On the other hand, reading comprehension section of UTEPT does not contain any DIF items with considerable effect size. In effect, it is the vocabulary section of UTEPT that is mainly responsible for academic discipline DIF.

All in all, detecting only four (4%) sizable DIF items out of a total of 100 items on UTEPT indicates that the test is not performing significantly differently for humanities and science and engineering. As a result, it can confidently be claimed that the size of academic discipline DIF is not big enough to render UTEPT unfair to examinees of humanities and science and engineering backgrounds.

There is a scarcity of research into content-based reasons of DIF in DIF literature (Uiterwijk & Vallen, 2005). This may partly be due to many uncertainties surrounding this type of research. First of all, the statistical indications of the presence of DIF in an item depend on the type of analysis: they may be present in all the DIF analyses, only in some of them, or in none at all. The fact is that the various item bias detection procedures do not produce the same results. Not every statistical procedure (for example, Mantel-Haenszel) can detect non-uniform DIF (Uiterwijk & Vallen, 2005). Secondly, it is not always clear which element in an item causes DIF.

Based on the content of DIF items with considerable magnitude, the content experts concluded that the sources of DIF in these items could be attributed to examinees' academic discipline. Item 33 as



the only item that favored science and engineering came from a cloze test on human biology. On the other hand, items 40, 50, and 52 which were significantly easier for humanities came from contents covering humanities topics. The sources of bias for all academic discipline items were hypothesized and were attributed to test takers' academic discipline. Thus, as the source of DIF is irrelevant to language ability of learners, that is the construct being measured, it can be concluded that we have a case of differential item validity, or item bias. Accordingly, it can be concluded that these items are biased against particular academic discipline groups and should be subject to further analysis in order to improve or abandon them.

It is to be remembered that, as it is also emphasized by Kunnan (1990), identifying potential sources of DIF is only the first step in DIF analysis. The next step is to determine what to do with items that display DIF and how to compensate test takers for such bias. Kunnan proposes two procedures to be used here: (a) sources causing DIF can be examined, hypothesized and causally related, if possible, and (b) items displaying DIF can be improved or discarded. Kunnan believes that "the second procedure should only be resorted to by test writers and administrators as a short term plan: to remedy a test in use, or when there are not enough resources for a detailed study" (1990, p. 6). The first procedure, he argues, which will have a more lasting effect, is recommended otherwise.

All in all, from the discussion of the DIF items above, it appears that it is sometimes very difficult to indicate with any certainty why one item shows significant DIF and another does not. That is why the researcher preferred to use terms such as '*possible*' and '*probable*' when attempting to hypothesize the sources of DIF.

Finally, it is important to note that fairness is a broad concept which includes much more than a mere DIF analysis of the items (Kunnan, 2010; Xi, 2010). Camilli (2006) also points out that DIF analysis that only looks at the performance of different *groups* rather than individuals cannot reveal the equally possible bias against particular individuals. In consequence, in order to assess the fairness of UTEPT, further research specifically focused on the performance of individuals rather than groups is needed.

## References

- Camilli, G., & Shepard, L. A. (1994). *Methods for identifying biased test items*. Thousand Oaks, SA: Sage.
- Clauser, B. E., & Mazor, K. M. (1998). Using statistical procedures to identify differentially functioning test items. *Educational Measurement*, 17(1), 31-44.



- Geranpayeh, A., & Kunnan, A. J. (2007). Differential item functioning in terms of age in the certificate in advanced English examination. *Language Assessment Quarterly*, 4, 190-222.
- Hale, G. (1988). Student major field and text content: Interactive effects on reading comprehension in the TOEFL. *Language Testing*, 5, 49–61.
- Holland, P. W. (1985). *On the study of differential item performance without IRT*. Paper presented at the meeting of the Military Testing Association, San Diego.
- Holland, P. W., & Thayer, D. T. (1988). Differential item functioning detection and the Mantel-Haenszel procedure. In H. Wainer & H. I. Braun (Eds.). *Test Validity* (pp. 129-145) Hillsdale, NJ: Erlbaum.
- Kim, M (2001). Detecting DIF across the different language groups in a speaking test. *Language Testing*, 18, 89-114.
- Kunnan, A. J. (1990). DIF in native language and gender groups in an ESL placement test. *TESOL Quarterly*, 24, 741-746.
- Liu, J., Harris, D. J., & Schmidt, A. (2007). Statistical procedures used in college admissions testing. In C. R. Rao & S. Sinharay (Vol. Eds.). *Handbook of statistics*, 26, 1057-1091.
- Penfield, R. D. (2009). *Differential item functioning analysis system*. DIFAS (5) user's manual.
- Rogers, H. J. & Swaminathan, H. (1993). A comparison of logistic regression and Mantel-Haenszel procedures for detecting differential item functioning. *Applied Psychological Measurement*, 17, 105-116.
- Swaminathan, H., & Rogers, H. J. (1990). Detecting differential item functioning using logistic regression procedures. *Journal of Educational Measurement*, 27, 361-370.
- Uiterwijk, H., & Vallen, T. (2005). Linguistic sources of item bias for second generation immigrants in Dutch test. *Language Testing*, 22, 211-234
- Zieky, M. (1993): Practical questions in the use of DIF statistics in test development. In P. W. Holland, & H. Wainer (Eds.). *Differential Item Functioning* (pp. 321-336). Hillsdale, NJ: Lawrence Erlbaum.
- Zumbo, B. D. (1999). *A handbook on the theory and methods of differential item functioning (DIF): Logistic regression modeling as a unitary framework for binary and Likert-type (ordinal) item scores*. Ottawa, ON: Directorate of Human Resource Research and Evaluation, Department of National Defense.

## Bio Data

**Seyyed Mohammad Reza Amirian** is assistant professor of TEFL at Hakim Sabzevari University. He got his PhD from the University of Tehran. He has published many articles in leading journals and presented in many national and international conferences. His research interest includes test fairness, DIF, test validation, and testing language skills.



## Assessment Practices: An Examination of Teachers' Perceptions of Classroom-Based Language Assessment

Masoomeh Estaji<sup>1</sup>

*Allameh Tabatabai University*

### Abstract

This study aims to investigate teachers' perceptions, beliefs, and their practice with respect to classroom-based English language assessment; hence it examines the teachers' current working principles of assessment and their practices. Four stages in teachers' classroom-based assessment were examined; planning, implementation, monitoring, and recording and dissemination. A questionnaire was developed reflecting these stages; its findings were analyzed statically and qualitatively. Further qualitative data was also collected and analyzed through interviews with volunteer participants. The results of the study revealed that the teachers hold and exercise their own firm beliefs regarding classroom-based assessment, and have a good knowledge of assessment or testing principles. It was found that teachers' beliefs in language learning, their understanding of learning goals, their preconceived ideas regarding students, and their prediction of the students' performance in the target language use domain influenced both their approach to planning assessment procedures and the actual implementation of assessment practices. Assessment "capacity" or "competence" has also been identified as a factor that affects teachers' assessment practices, and the importance of language teachers' professional training and development in assessment has been raised. However, it became evident that the teachers were constantly developing their skills and knowledge regarding assessment in order to address any possible challenges or tasks given to them. Finally, certain areas needing further investigation were identified.

**Keywords:** Assessment, Classroom-based language assessment, Teachers' perception of assessment, Teachers' assessment practices

### 1. Introduction

Classroom assessment is a major component of classroom research, which involves teachers in the continuous monitoring of students learning (Angelo & Cross, 1993). Furthermore, Angelo and Cross (1993) state that as feedback, classroom assessment provides teachers with feedback about their effectiveness as teachers and it gives students a measure of their progress as learners. In some ways, classroom assessment helps teachers obtain useful feedback on what, how much, and how

---

<sup>1</sup> Email: mestaji74@gmail.com



well their students are learning. Teachers can then use this information to refocus their teaching to help students make their learning more efficient and more effective. Moreover, through practice in classroom assessment, teachers are able to better understand and promote learning, and increase their ability to help students themselves become more effective, self-assessing, and self-directed learners. Simply put, the central purposes of classroom assessment are to empower both teachers and their students to improve the quality of learning in the classroom (Angelo & Cross, 1993; Cohen, 1994).

Therefore, this study focuses on the teacher as “the agent of the assessment process” (Harlen, 1996, p. 129; Rea-Dickens, 2004, p. 251) because teachers’ beliefs and attitudes have a great influence on their assessment and evaluation practices (Arkoudis & O’Loughlin, 2004; Breen et al., 1997; Davison, 2004; Edelenbos & Kubanek-German, 2004). This study might also contribute to an emerging interest in teachers’ classroom-based assessment and their role as assessors in the field of applied linguistics and English teaching. It is these considerations that have led this research to focus on the context in which teachers operate, and their specific understandings or perceptions of classroom-based English language assessment.

## 2. Theoretical Background

Education is said to change learners’ behavior desirably, and the quantity and quality of such changes are determined by assessment. Further, with the increased focus of society and educational agencies on standardized test scores, great scrutiny has been placed on assessment. Recent years have seen increased research on classroom assessment as an essential aspect of effective teaching and learning (Bryant & Driscoll, 1998; McMillan, Myran & Workman, 2002; Stiggins, 2002). It is becoming more and more evident that classroom assessment is an integral component of the teaching and learning process (Gipps, 1990; Black & Wiliam, 1998). According to Dufresne, Gerace, Mestra, and Leonard (2000) assessment informs the teacher about what students think and about how they think. Classroom assessment helps teachers to establish what students already know and what they need to learn.

However, teachers and program administrator are sometimes in struggle to identify appropriate procedures to assess the knowledge and abilities of English language learners (O’Malley & Pierce, 1996, p. 3). Difficulties that they face are due to the need to identify varying levels of knowledge and proficiency in English and various or complex purposes the assessment possess.

With reference to teachers’ roles as assessors, Davison (2004, p. 325) suggests a cline of teachers as assessors as determined by their beliefs and practices. However, there might be tension



between the different roles of teachers because they find themselves at the confluence of different assessment practices, such as playing a role “as facilitator and monitor of language development, and that of assessor and judge of language performance as achievement” (Rea-Dickens, 2004, p. 253). Taking this into account, language teachers should be aware of who they are and what they are in the language assessment process for. Consequently, it is important to consider teachers’ pre-existing beliefs about the students and the targets and material they are assessing, and the social, institutional and cultural context of their assessment environments. With regard to this, Rea-Dickins and Rixon (1999) argue that what is required of teachers if they are to be regarded as good assessors is that: To begin with, teachers need to know more about the theories of assessment and testing in order to engage in the process; this requires them to understand the assessment culture in which they are working because it might be different across districts or countries. In addition, they should have access to information about any testing and assessment they are supposed to carry out. Finally, more significantly, in case they are asked to be involved in the marking of students’ performance or to design their own assessment procedures, they should receive relevant training.

For the above reasons, it is essential to investigate the ‘teachers’ perceptions of themselves as assessors in order to understand the reality of teachers’ classroom-based assessment, and to provide them with such teacher development programs as are needed in order to help them enhance their expertise in teaching and assessment. It might be expected that through this process teachers would have an opportunity to reflect for themselves on why they do what they do when assessing their students.

Hence, the way teachers perceive assessment may influence the way they teach and assess their students (Assessment Reform Group, 1999; Fennema & Romberg, 1999). This study investigated the current classroom assessment practices of teachers in two selected institutions in Iran. To accomplish the purpose of this study, the following research questions were developed:

1. What is the influence of teachers’ perceptions of classroom assessment on the ELT teachers and their classroom assessment practices?
2. How do teachers of high-proficiency levels (i.e., Intermediate and Advanced) perceive classroom assessment?



### 3. Method

#### *3.1. Participants*

Thirty six Intermediate and Advanced English teachers drawn from two language institutions in Iran participated in this study. Eighteen teachers were drawn from each of the two language institutions. All the participants were female and their teaching experiences ranged from 3 to 17 years. The participants were trained teachers currently teaching high-proficiency levels in the two institutions. All potential participants had had at least three years of teaching experience at the institution and as a consequence had made assessment-related decisions for the students that ranged from grading individual assignments to marking final exams and making end-of-year pass/fail decisions. Of these participants, 5 teachers had a doctorate in English language teaching (ELT); 16 had Master's degrees in ELT; 15 were reading for a Master's degree in ELT in Iran. In addition, all the teachers had sufficient experience of, and interest in English language teaching to make them suitable to participate in this study. Teachers were selected based on gender, classes they were teaching, and teaching experience. The trend at both institutions was not very different from each other.

#### *3.2. Instrumentation*

In order to address the research questions, 2 instruments were used: questionnaire surveys and interviews were conducted in order to obtain the teachers' ideas, opinions, and experiences with regard to their classroom-based assessment; assessment batteries used were also collected and analyzed as they showed what the teachers had done in the classroom.

#### *3.3. Data collection procedure*

The initial data collection process included the pilot testing of the questionnaire and interview protocol. Teachers with similar characteristics to those who participated in the actual study were sampled for the pilot phase of the study. Some questions on the questionnaire and interview protocol were modified after feedback from the teachers. Those who participated in the pilot phase were not allowed to participate in the actual data collection phase.

Data were collected within eight weeks that included both quantitative and qualitative data. A written questionnaire was distributed to the teachers in the institution. This included sets of closed questions and concluded with an option to comment on any uncovered issues. Teachers were assured of anonymity and confidentiality as their data was only to be seen by the researcher. The teachers were also interviewed using audio recording. The scope of the interview was deliberately



broad and initially semi-structured in order to provide a framework for the participant to tell their own story, express their understanding, and give their impression of their language assessment in their classrooms.

#### **4. Data Analysis**

Quantitative and qualitative data were analyzed separately given different nature of the data; however, after analysis the data were integrated to support themes and various points. In order to estimate if the gap between teachers' beliefs and their practice is statistically significant, a t-test was executed. Moreover, the aims of the interview were to investigate the gaps between the teachers' beliefs and their practice, and to gain some understanding of the reasons for perceived divergences between principles and practice in their assessment activities. All the interviews were recorded by voice recorder and saved in an audio file in MP3 format. Then they were translated and analyzed through classification and listing. That is, the teachers' responses were classified based on such criteria as: whether or not they represent the reality of teachers' classroom-based assessment in a given context; whether or not they might contribute to the development of the body of theory of classroom-based assessment; and whether or not they might provide the teachers with any clues as to how to improve their assessment practices.

#### **5. Results**

##### *5.1. Quantitative results: The gaps between the teachers' working principles and their practice*

A paired sample t-test was conducted to compare the group means in terms of their perceptions of classroom-based assessment and their practices reflecting the working principles of classroom-based assessment. The results revealed that the teachers' perception of classroom-based assessment has a great impact on their practices. However, it also became clear that these teachers did not put some of their principles into practice. Table 6 reveals the results of the paired sample t-test in the planning stage.



**Table 6.**

Paired Sample T-test of the Planning Stage

	Paired Differences					t	df	Sig. (2-tailed)		
	Mean	Std. Deviation	Std. Error	95% Confidence Interval of the Difference						
				Mean	Lower	Upper				
Pair 1 Q1-Q1'	.198	.570	.061	.075	.320	3.215	35	.002		
Pair 2 Q2-Q2'	.058	.561	.060	-.062	.178	.962	35	.339		
Pair 3 Q3-Q3'	.465	.877	.095	.277	.653	4.918	35	.000		
Pair 4 Q4-Q4'	.547	.762	.082	.383	.710	6.654	35	.000		
Pair 5 Q5-Q5'	.384	1.065	.115	.155	.612	3.343	35	.001		
Pair 6 Q6-Q6'	.151	.584	.063	.026	.276	2.399	35	.019		
Pair 7 Q7-Q7'	.477	.822	.089	.300	.653	5.378	35	.000		
Pair 8 Q8-Q8'	.151	.744	.080	-.008	.311	1.884	35	.063		
Pair 9 Q9-Q9'	.174	.770	.083	.009	.339	2.101	35	.039		
Pair 10 Q10-Q10'	.465	.864	.093	.280	.650	4.994	35	.000		
Pair 11 Q11-Q11'	1.000	1.117	.120	.761	1.239	8.304	35	.000		
Pair 12 Q12-Q12'	.512	.851	.092	.329	.694	5.578	35	.000		
Pair 13 Q13-Q13'	-.105	.933	.101	-.305	.095	-1.040	35	.301		
Pair 14Q14-Q14'	.756	1.227	.132	.493	1.019	5.715	35	.000		
Pair 15 Q15-Q15'	-.047	.969	.104	-.254	.161	-.445	35	.657		
Pair 16 Q16-Q16'	.221	.726	.078	.065	.377	2.822	35	.006		
Pair 17 Q17-Q17'	.209	.922	.099	.012	.407	2.106	35	.038		

There were many items which showed that the differences between teachers' beliefs and their practice was statistically significant; but, as statistical significance just indicates there were some items on which the teachers had similar views as their practices in the classroom.

Alternatively, the statistical results of the paired sample t-test revealed a similar trend for the implementation, monitoring, and reporting and dissemination stages of classroom assessment.



### *5.2. Qualitative results: Teachers' perception of classroom assessment*

Definitions provided by the respondents answer part of the first research question; teachers' perceptions of classroom assessment. Definitions given by the teachers will shed light on how they see classroom assessment based on their position as language teachers. Various definitions were presented by the teachers regarding the *what & why* of classroom assessment. They expressed their view points by discussing the forms and purposefulness of assessment in different classroom environment. According to an English teacher, "assessment is what we monitor students' academic and psychological activities, we observe where they started and where they are now". In addition, other teachers agreed that the goal of assessment depends on the time assessment is implemented that occurs daily and weekly to the entire semester period. They made a distinction between two forms of assessment, formal and informal assessment.

According to a language teacher, "formal assessment is another word for measurement that we use in mid-term and final exam, checking students' mastery of the content"; informal assessment, on the other hand, she explained, occurs during the instruction "with an emphasis on inclusion of assessment in the daily instruction". The distinctions that the teachers make between assessment forms, imply the extent the teachers are exposed to this concept and with that in mind, their assertion whether they are applying those methods in their current practices of assessment can be revealed.

In addition, another English instructor who enjoyed more than eight years of instruction experience believed that "assessment is a broad concept that cannot be summarized only during the instruction, it has two forms: introductory assessment and periodic assessment that occurs during the instruction". However, the majority of instructors had little emphasis on the formative side of assessment; they were often concerned about the achievement part that occurs at the end. Linn and Miller (2005) characterized them as assessment for learning and assessment of learning with an emphasis on the importance of the first one that occurs during the instruction.

Based on the distinction made, a well-experienced instructor described assessment as "tests that teachers give students to make sure students learned something and the teacher could transfer the lesson". Similarly, another instructor supported, "assessment is a mutual activity between students and the teacher". However, one another instructor, among the respondents had a different view in terms' of classroom assessment. He labeled assessment as a sub-part of an exam, "assessment is a small section of evaluation," she explained.



To sum up, an analysis of the overall perspectives of the teachers in terms of defining the *what* and *why* of assessment shows that teachers had a recognition of various forms and purposes of classroom assessment. In addition, they viewed assessment as activities for educational purposes not just assigning score or grade, although to some extent teachers leaned toward the achievement unit that is assessed at the end of a course. Additionally, some level of disagreement existed among the teachers in terms of the scope of assessment: some viewed assessment encompassing all activities during the course, while a few who didn't enjoy a long-term of teaching experience viewed it as a small segment of an exam, which implies that not all the instructors had been exposed to the concept assessment. Overall, responses from the teacher respondents attest that they had a significant voice regarding the definition and the purpose of classroom assessment.

## 6. Discussion

This study has tackled the issue of the teacher as an "agent in the assessment process" (Rea-Dickens 2004, p. 252) and as an assessment developer in the area of English language assessment. The study found that teachers recognized what good assessment is, that they usually had their own assessment principles, and put these into practice. They were also able to base classroom-assessment on their own professional expertise as language teachers cum assessors; this was the case when they faced matters which needed to be addressed during the progression of the assessment cycle. They tried to consider all the possible factors which should be taken into account when assessing their particular students. Thus, teachers' classroom-based assessment might be recognized and could be developed as one of the distinctive and independent areas of language assessment. Therefore, there is a need for a considerable amount of in-depth investigation into teachers' convictions and their practice; this might include the following themes: teachers' diagnostic competence' (Edelenbos & Kubanek-German, 2004) and their actual diagnostic activities in assessment activities; a more detailed examination of the decision making processes; teachers' decision making process as reflective practitioners especially when they face the conflict between their beliefs and the principles or rules of assessment which are imposed upon them by the authorities and other outside influences; the strategy they adopt for giving feedback based on the assessment process; the process or strategy for the development of the assessment specifications and tasks; and finally, their strategy for developing themselves into better teachers and assessors of language. The core rationale for all these is that it is the teachers' beliefs and their attitudes that have a significant influence on their assessment and evaluation practices (Arkoudis & O'Loughlin, 2004; Breen et al., 1997; Davison, 2004). Furthermore, there is a need for substantial research on the policies governing English teaching and assessment and on how these might be positively or negatively affecting what has happened in English education, whether consciously or unconsciously.



Taking into account the fact that students' everyday school lives and their academic results are hugely influenced by their classroom teachers, it is clear that there must be close cooperation between the classroom teachers and teachers of English; it is essential to enhance the students' progress in English. Thus, it might be wise to clarify differences between the roles of the classroom teachers and the teachers of English; an interchange of ideas concerning the process of English teaching and assessment should follow. To ensure that this takes place, a specific action plan should be designed to address the issue. The ultimate aim is to ensure that everything possible is done to encourage and support the students as they study English, and that both English specialists and classroom teachers play their part in achieving this.

Assessment "capacity" or "competence" has also been identified as a factor that affects teachers' assessment practices (Cheng, Rogers, & Hu, 2004), and the importance of language teachers' professional training and development in assessment has been raised (Bachman, 2000; Brindley, 1998; Edelenbos & Kubanek-German, 2004). This also relates to Rea-Dickens's (2008) point of view that language teachers should have a firm understanding and awareness of both language and the developmental processes involved in language acquisition. Given that general educational assessment knowledge and skills are documented (Standards for Teacher Competence in Educational Assessment of Students, 1990), it would be largely helpful if guidelines necessary for teacher competence pertaining to language teaching were available. Continued interest in the array of research that seeks to further identify and depict teachers' perception, cognition, and competence, as well as factors that influence their decisions when using assessment in the classroom, is conducive.

## References

- Angelo, T. A., & Cross, P. K. (1993). *Classroom assessment techniques* (2nd ed.). San Francisco: Jossey-Bass.
- Arkoudis, S., & O'Loughlin, K. (2004). Tensions between validity and outcomes: teacher assessment of written work for recently arrived immigrant ESL students. *Language Testing*, 21(3), 284-304.
- Assessment Reform Group (1999). *Assessment for learning*. Retrieved on 30 August, 2004 from: [http://www.aaia.org.uk/pdf/Assessment\\_for\\_Learning.pdf](http://www.aaia.org.uk/pdf/Assessment_for_Learning.pdf)
- Bachman, L. (2000). Modern language testing at the turn of the century: Assuring that what we count counts, *Language Testing*, 17(1), 1-42.
- Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education*, 5(1), 7-74.



- Breen, M. P., Barratt-Pugh, C., Derewianka, B., House, H., Hudson, C., Lumley, T., & Rohl, M. (1997). *How teachers interpret and use national and state assessment frameworks*. Canberra: Department of Employment, Education, Training and Youth Affairs.
- Brindley, G. (1998). Outcomes-based assessment and reporting in language learning programs: A review of the issues. *Language Testing*, 15(1), 45-85.
- Bryant, D., & Driscoll, M. (1998). *Exploring classroom assessment in mathematics: A guide for professional development*. National Council of Teachers of Mathematics: Reston.
- Cheng, L. Rogers, T. & Hu, H. (2004). ESL/EFL instructors' classroom assessment practices: Purposes, methods, and procedures. *Language Testing*, 21(3), 360-389.
- Cohen, A. D. (1994). *Assessing language ability in the classroom* (2nd ed.). Boston, MA: Heinle and Heinle.
- Davison, C. (2004). The contradictory culture of teacher-based assessment: ESL teacher assessment practices in Australian and Hong Kong secondary schools. *Language Testing*, 21(3), 305-334.
- Dufresne, R. J., Gerace, W. J., Mestre, J. P., & Leonard, W. J. (2000). *ASK/IT /A2L: Assessing student knowledge with instructional technology* (No. PERG-2000#09-Sep#1-28pp). Amherst, MA: University of Massachusetts Physics Education Research Group.
- Edelenbos, P., & Kubanek-German, A. (2004). Teacher assessment: The concept of diagnostic competence. *Language Testing*, 21(3), 259-283.
- Fennema, E., & Romberg, T. A. (1999). *Mathematics classrooms that promote understanding*. New Jersey: Lawrence Erlbaum Associates.
- Gipps, C. (1990). *Assessment: A teachers' guide to the issues*, London: Hodder and Stoughton.
- Harlen, W. (1996). Editorial. *Curriculum Journal*, 7, 129.
- Linn, R. L., & Miller, M. D., (2005). *Measurement and assessment in teaching*, (9th ed.). New Jersey: Pearson.
- McMillan, J. H., Myran, S., & Workman, D. (2002). Elementary teachers' classroom assessment and grading practices. *Journal of Educational Research*, 95, 203-213.
- O'Malley, J. M. & Pierce, L.V. (1996). *Authentic assessment for English language learners*. New York: Addison-Wesley.
- Rea-Dickens, P. (2004). Understanding teachers as agents of assessment. *Language Testing*, 21(3), 249-258.
- Rea-Dickens, P. (2008). Classroom-based language assessment. In E. Shohamy & N. H. Hornberger (Eds.), *Encyclopedia of language and education*, (2nd ed.), Vol. 7. *Language Testing and Assessment*, (pp. 257-271). Heidelberg, Germany: Springer.
- Rea-Dickins, P., & Rixon, S. (1999). Assessment of young learners' English: Reasons and means. In S. Rixon (Ed.), *Young learners of English: Some research perspectives* (pp. 89-101). Harlow, Essex: Longman.



Standards for teacher competence in educational assessments of students (1990). Washington, DC: American Federation of Teachers, National Council on Measurement in Education, and National Educational Association. Retrieved March, 2011 online at: <http://www.unl.edu/buros/bimm/html/article3.html>

Stiggins, R. J. (2002). Where is our assessment future and how can we get there from here? In R. W. Lissitz & W. D. Schafer (Eds.). *Assessment in educational reform: Both means and ends* (pp. 112-125). Boston: Allyn and Bacon.

### Bio Data

**Masoomeh Estaji** is assistant professor of TEFL at Allameh Tabatabai University. She holds a Ph.D in TEFL from Allameh Tabatabai University. She has participated in numerous local, national workshops on educational issues, methodology, and testing. She has also presented numerous papers on methodology, testing, and education at a number of TEFL conferences. Her research interests include Language Testing and Assessment, SLA and Psycholinguistics.



## The Role of Resource Management Strategies on the English Language Learners' Achievement

**Maryam Mohammadi<sup>1</sup>**

*Islamic Azad University, Ardabil Branch*

**Afsar Rouhi**

*Payam-e-Noor University, Tehran Branch*

**Mehran Davaribina**

*Islamic Azad University, Ardabil Branch*

### **Abstract**

The purpose of this study was to investigate the relationship between students' resource management strategies and their academic achievement. The research method was descriptive-correlation. Statistical population was about 300 female and 200 male students in English literature majors from Islamic Azad University, Ardabil Branch who were freshmen and seniors. The sample included 81 females and 56 males who were selected randomly via Krejcie and Morgan Table considering alpha = .05. The instrument of this research was the resource management strategies dimension of motivated strategies for learning questionnaire. Reliability of questionnaire and its subscales was calculated using Cronbach's Alpha. Also academic achievement was obtained through asking a final grade average from the students. Gathered data were analyzed using one sample *t*-test, and regression. A regression analysis revealed that the resource management strategies subscales including time and study environment management, effort regulation, peer learning, and help seeking explained 36 percent of variance for achievement of English learners ( $R = .60$ ,  $R^2 = .36$ , adjusted  $R^2 = .33$  and  $F_{(4,132)} = 18.05$ ;  $p < .01$ ).

**Keywords:** English as a foreign language, Resource management strategies, Academic achievement, Effort regulation

### **1. Introduction**

In today's Information Age, English language has increasingly dominated international communication and information access. So, the use of English as a common language is not controversial (Canagarajah, 2002; Tardy, 2004). So, it could be said that the learning English is vital especially for whom English is a foreign language. As Williams and Burden (2000) stated learning a

---

<sup>1</sup> Email: M\_maryam5980@yahoo.com



foreign language is different to learning other subjects, mainly because of the social nature of such a venture.

The results of some studies (Abbasi, Ahmadi, & Lotfi, 2009; Bigdeli, 2010) and experiments of researcher and statements of teachers indicated that there were difficulty in teaching and learning of English language in our country. Many of students hadn't good remembrance of learning English.

In the area of English learning many experts agreed that in the same class there were students who differ from each other in learning English language. So, nowadays focus of instructional experts was shifting from the teacher-centered classroom to a more learner-centered classroom (Sheikholeslami & Khayer, 2006). Sadighi and Zarafshan (2006) stated that the goal of teachers and trainers was shifting the focus of learners as dependent learner to independent and autonomous learner in learning.

Learners need to manage learning resources along with learning strategies. Learning resource management strategies are facilitating strategies and learners usually use them to control and to administer the environment, time, effort, and help-seeking from people of the same age or others (Pintrich, 2004). These strategies help learners to control available resources in order to become compatible with the task. In other words, it is under the shadow of learning resource management strategies; learners gain the ability to do their tasks (Kivinen, 2003). They also demonstrate the quality and quantity of learners' involvement with the task and it includes planning in line with the set objectives in order to gain better results (Filcher & Miller, 2000).

The findings of Campbell (2007) demonstrated that academic performance has positive and significant relationship with the effective application of resource management strategies in a way that lack of mastery of learners on learning resource management strategies will create many problems for them and they will usually experience academic failure and will become disappointed. In theory of Pintrich (2003) learning resource management strategies include: time management, studying environment, effort regulation, learning from people of the same age and help-seeking.

### *1.1. Time management and studying environment*

Time management is today one of the determining factors of success and it has gained importance (Arnold & Pulich, 2004). In way that some experts believe that avoiding waste of time and control of time can increase learners' capabilities, reduce stress, and increase their mental health (Gran-Moravec & Hughes, 2005). Therefore, Koszin (2007) stated that learners need to be able to manage their studying time. Also Campbell (2007) said that time management includes planning, design, and the individual's studying time. Researchers have demonstrated that design and



management of time helps learners to be more self-regulatory in their use of studying time and thus academic achievement.

Britton and Tessor (1991) found out that there is a positive and meaningful relationship between students' belief in time management and their progress in the short term. Most people imagine that they cannot control time and they cannot save or move on their will. However, time management is in actual fact the life management, individual management, and self-management. According to Vecchio (2003), time management means self-management and those who manage their time can do tasks more effectively and more consciously. Terry (2002) reports a positive relationship between time management, self-regulation, and self-efficacy.

In addition to time management, environment management is also of great importance and it needs to be placed somewhere that is calm and is almost free of all audio and visual disturbances in a way that guarantees the individual's concentration (Chen, 2002). In Ford's theory of motivational systems, environment management is one of the main factors of the responding situation and the individual needs to choose an organized, calm and almost free of interruptions when doing a task (Campbell, 2007). However, researches are of the conviction that regulation of studying environment is based on the learner's learning style in a way that he feels that this environment will be the proper grounds for his learning (Filcher & Miller, 2000).

According to Purdie and Hattie (1996), Zimmerman and Martinez-Pons (1986), self-regulated learners initiate efforts to organize the learning context in ways that help them to learn better. They say structuring of their learning environments involves two forms. The first form is the deliberate selection of a place in which to work or manage a setting to make learning easier (e.g., "I isolate myself from anything that distracts me."). The second form is the performance of a particular personal behavior so that learning is improved (e.g., "I carefully investigate and fully utilize the resources of the library.); "When preparing for a class presentation, I try to anticipate any questions the audience might have.").

### *1.2. Effort regulation*

Kosnina (2007) indicated that effort regulation or willingness to have concentration and effort to gain the goals occurs in spite of potential interruptions. He says in academic situations, effort regulation affects the person's learning style and enables the learners to control and manipulate some outside system interruptions. Effort management or effort regulation is a process in which learners use such techniques as effort, self-tell, perseverance and self-encouragement (Filcher & Miller, 2000).



### *1.3. Help-seeking from classmates*

Salmons (2008) has categorized the co-operative levels between learners into five groups on the basis of trust. The lower levels of co-operation need less trust (whether organizational or personal) and the higher levels of co-operation need higher trust. This trust is to be between the learners themselves and between the learners and the institute and teachers. The five levels of co-operation, in ascending order, are:

- Discussion: is the first level in Salmons category. Learners coordinate between seemingly separate ideas and projects and have a regular and organized effort to create shared knowledge.
- Criticism and analysis of peers: is the second level of co-operation and describes a critical and feedback-based process among the participants. In this level, learners judge the quality and amount of information relationship presented by their classmates. In the shadow of this judgment, participants learn a lot from one another.
- Parallel co-operation: is the third level of co-operation in which each part of the task is given to one of the members. Members usually work individually and combine the results of their efforts through discussions and criticism and analysis of their peers in order to present one single outcome.
- Sequential co-operation: is the fourth level of Salmons levels of co-operation. In this level, a group attempts to use a sequential structure to do a task. On this basis, the different components of tasks are divided into sequential steps and the results obtained in each stage will be merged as the final outcome. In this process each step depends on the successful completion of the previous steps and it is merged at each step after completion by individual attempts and through the process of discussion and peer analysis.
- Additional co-operation: is the last level of co-operation in which a group of cooperative learners use an additive structure and cooperate with each other at all stages and merge their ideas and projects for planning, organizing, and performing the task.

### *1.4. Help-seeking*

Some researchers such as Rickwood, Deane, Wilsons, and Ciarrochi (2005) indicated that help-seeking is one of the efficient ways of combating problems and it results in the reduction of psychological confusion and increase of learning strategies. It is an effective way of managing learning when students have enough understanding or knowledge for dealing with an issue on themselves. This characteristic not only helps learners to cope with their learning needs



immediately, but is an effective way to enhance and improve their performance. Researches demonstrated that some learners have help-seeking skill for improving their learning, while others lack this ability (Taplin, Yum, Jegede, Fan, & Chan, 1999). The learners who believe in self-autonomy (Wilson & Deane, 2002) have a negative attitude toward help-seeking (Taplin et al., 1999) rarely ask for help.

Help-seeking, as one way of learning management strategy, is affected by social motivations in contrast to other learning strategies in which there is social interaction (Chen, 2002). Pakdaman, Farzad, Sarmad, and Khanzadeh (2007) believed that help-seeking is affected by cultural beliefs and values. They noted that in the societies, which emphasize the autonomy, self-reliance, and individuality, help-seeking is considered an independent behavior. In contrast, the societies have a community-based attitude; help-seeking is a social interaction and compatibility behavior. However, mere help-seeking cannot be the criteria of judgment about compatibility or non-compatibility and not all kinds of help-seeking lead to progress.

Rickwood, et al. (2005) noted that help-seeking is the process of identification of need, making decision to seek help, and performing the decision and it is not easy. They argued that in the practical situations, each stage is interrupted by different factors. Help-seeking can possibly not be recognized as a need, or if so, it might not be realized in the form of intention or demand, and it eventually might not lead to a certain behavior.

Karabenick and Knapp (1991); Newman and Schwager (1993) and Zimmerman and Martinez-Pons, (1986) found that based on self-regulation theory when self-regulated learners faced with difficult tasks that require assistance from a more knowledgeable person, self-regulated learners seek help from their instructors or classmates. Schunk and Zimmerman (1994) stated that help-seekers can "act purposefully and instrumentally, not only remedying an immediate problem but ensuring long-term autonomy through mastery of a task" (p. 284). If they do not know an answer or do not understand a lesson, they ask questions (e.g., "If I do not understand something during a class meeting, I will ask for additional clarification"). Considering all pertinent and available information, they make the following decisions: (1) Is it necessary that I ask for help? (2) What should I ask? (3) Whom should I ask (the instructor or a classmate)? And (4) What is the most suitable way of requesting help in this particular circumstance? Briefly, environmental structuring and help-seeking are viewed as ways of achieving environmental control and utilization in the learning process. These play an important role in self-regulated learning.



## 2. Method

### 2.1. Participants

Statistical population of this research was about 300 female and 200 male students in English literature majors from Islamic Azad University, Ardabil Branch who were freshmen and seniors. The sample included 81 (59.1 %) females and 56 (40.9 %) males who were selected by using combination of random and cluster. The alpha level was set at .05. The number of sample was calculated via Krejcie and Morgan Table.

### 2.2. Instrument

The instrument of this research was the Resource management strategies dimension of motivated strategies for learning questionnaire (MSLQ) instrument. Table 1 illustrates the Resource management strategies including 4 subscales. Items were scored on a 7-point Likert scale, from 1 (not at all true) to 7 (very true).

**Table 1.**

Coefficient alphas and 4 subscales of the resource management strategies

Scale	Items	Reliability*	Reliability**
Time and Study Environment Management	35, 43, 52 <sup>r</sup> , 65, 70, 73, 77 <sup>r</sup> , 80 <sup>r</sup>	.76	.77
Effort Regulation	37 <sup>r</sup> , 48, 60 <sup>r</sup> , 74	.69	.71
Peer Learning	34, 45, 50	.76	.77
Help Seeking	40 <sup>r</sup> , 58, 68, 75	.52	.71
Resource management strategies	81	-	.70

*r: revised scoring    'Source: (Duncan & McKeachie, 2005); \*\* present study*

The reliability of questionnaire and its subscales was calculated using Cronbach's Alpha (Table 1). The calculated values showed that the questionnaire and subscales of questionnaire enjoyed high internal consistency. Also, the rate of academic achievement was obtained through asking the students on their final grade average. The mean score of each component and its subscales were compared using one sample *t*-test with conceptual mean score. The conceptual mean score was considered 4, because the questionnaire was 7-point Likert scale whose median was 4. The normality of the data distribution was examined through Kolmogorov-Smirnov test.



**Table 2.**

One-sample kolmogorov-smirnov test for normality

Scale	Most Extreme Differences			K-S z	Sig.	Distribution
	Absolute	Positive	Negative			
Time and Study Environment Management	.066	.055	-.066	.768	.597	Normal
Effort Regulation	.108	.108	-.065	1.26	.084	Normal
Peer Learning	.088	.058	-.088	1.03	.243	Normal
Help Seeking	.094	.048	-.094	1.10	.179	Normal
Resource management strategies	.044	.044	-.038	.516	.953	Normal
Academic achievement	.080	.080	-.057	.933	.349	Normal

*N* = 137

As it is clear in Table 2, amounts of Kolmogorov-Smirnov Z calculated based on  $p > .05$  indicated that distribution of the total variables was normal.

### 3. Results

*RQ: What is the rate of resource management component among female and male students?*

Table 3 shows that the mean scores of time and study environment management for female students ( $4.0 \pm .84$ ) and male students ( $3.8 \pm .77$ ). While the rate of females' time and study environment management strategy were equal with median and males' time and study environment management strategy were statistically lower than the median (4), but calculated  $t_{(80)} = .43$ , ( $p = .669$ );  $t_{(55)} = -1.5$ , ( $p = .135$ ) showed that the difference were not statistically significant in female and male students' in time and study environment management.

The mean scores of effort regulation in female students ( $4.5 \pm 1.2$ ) and male students ( $4.3 \pm .93$ ) were statistically higher than the mean scores (4). Based on *t*-test analysis  $t_{(80)} = 3.4$  and  $t_{(55)} = 2.1$  for female and male it became apparent that the two groups in effort regulation were higher than the median and the difference was statistically significant ( $p < .05$ ). The mean scores of peer learning for female students ( $3.9 \pm 1.5$ ) and male students ( $4.1 \pm 1.33$ ) were calculated. While the rate of peer learning strategy for females were statistically lower than the median and males' peer learning strategy were statistically higher than the median (4), but computed  $t_{(80)} = -.46$ , ( $p = .643$ );  $t_{(55)} = .60$ , ( $p = .550$ ) showed that the difference in two groups were not statistically significant. Finally, the mean scores of help seeking for female students ( $4.6 \pm 1.2$ ) and male students ( $4.5 \pm 1.1$ ) were



statistically higher than the mean scores (4). Calculated  $t_{(80)} = 4.5$  and  $t_{(55)} = 3.7$  showed that the results were statistically higher than the median ( $p < .05$ ).

**Table 3.**

Rate of resource management component and its subscales among female and male students

Variable	Gender	N	Mean	Std. Deviation	t	df	Sig.
Time and study environment management	Female	81	4.0	.84	.43	80	.669
	Male	56	3.8	.77			
Effort regulation	Female	81	4.5	1.2	3.4	80	.001
	Male	56	4.3	.93			
Peer learning	Female	81	3.9	1.5	-.46	80	.643
	Male	56	4.1	1.33			
Help seeking	Female	81	4.6	1.2	4.5	80	.000
	Male	56	4.5	1.1			
Resource management component	Female	81	4.2	.80	2.7	80	.012
	Male	56	4.1	.80			

Test value = 4

The mean scores of resource management component for female students ( $4.2 \pm .80$ ) and male students ( $4.1 \pm .80$ ) were statistically higher than the mean scores (4). The t-test for this issue revealed that female students' resource management component was statistically higher than the median and the difference was statistically significant ( $t_{(80)} = 2.7$ ,  $p < .05$ ), but male students' resource management component was not statistically higher than the median ( $t_{(55)} = 1.1$ ,  $p = .259$ ).

*RH: There is a relationship between resource management subscales with academic achievement.*

As can be seen from Table 4,  $R = .60$ ,  $R^2 = .36$ , adjusted  $R^2 = .33$  and  $F_{(4,132)} = 18.05$  were statistically significant ( $p < .01$ ). So it can be concluded that the resource management strategies subscales including time and study environment management, effort regulation, peer learning, and help seeking predict students' language learning significantly. In other words, 35 percent of variance for achievement of English learners could be explained via time and study environment management, effort regulation, peer learning, and help seeking. In the subscales of resource management strategies, the shares of effort regulation, and help seeking ( $Beta = .378$ , and  $Beta =$



.175, respectively) for predicting academic achievement were statistically significant. But the shares of time and study environment management, peer learning (*Beta* = .081, *Beta* = .124, respectively) were not statistically significant.

**Table 4.**

Predicting the academic achievement based on the subscales of resource management strategies

Model	S.S.	Df	M. S.	F	Sig.
Regression	203.9	4	50.98	18.05	.000
Residual	372.8	132	2.82		
Total	576.7	136			
<b>Coefficients</b>					
Model	Unstandardized coefficients		Standardized coefficients	t	Sig.
	B	Std. Error	Beta		
Constant	8.91	.81		11.03	.000
Time and study environment management	.025	.03	.081	.96	.341
Effort regulation	.178	.04	.378	4.52	.000
Peer learning	.059	.04	.124	1.43	.155
Help seeking	.077	.04	.175	2.03	.044

R = .60, R<sup>2</sup> = .36, Adj. R<sup>2</sup> = .33

Predictors: (constant), time and study environment management, effort regulation, peer learning, help seeking

Dependent Variable: English achievement

#### 4. Discussion

The research question was related to the rate of resource management and subscales between females and males. Findings indicated that help seeking (female), help seeking (male), effort regulation (female), peer learning (male) and effort regulation (male) were higher than the median. But peer learning (female), time and study environment management (female and male) were not higher than the median. The studies done by researchers such as Life (2011) indicated that English students prefer group learning; they like to learn English inside and outside of class.



A regression analysis revealed that the resource management strategies subscales including time and study environment management, effort regulation, peer learning, and help seeking explained 36 percent of variance for achievement of English learners. The shares of effort regulation, and help seeking were statistically significant. But the shares of time and study environment management, peer learning were not statistically significant.

The significance of the relationship between learning resources management strategies and academic achievement can be specified in the study of Pintrich (2004). He believes that learning resource management strategies are strategies for facilitating the learning and controlling and managing the environmental factors. Also Kivinin (2003) emphasis these strategies as learners' assistance in controlling resources available to them, Filcher and Miller (2000) propos these are effective factors in engagement of learners in tasks, and also Campbell (2007) believes that incompetency of learners in learning resource management skills leads them to serious difficulties in learning procedure.

The significance of the simple relationship between subscales of learning resources management and students' academic achievement has been emphasized and supported in most of the theories and researches, as Gran-Moravec and Hughes (2005); Kosnin (2007) believed that time management can enhance the ability and potential for employing capabilities, reducing stress, and increasing mental health. Campbell (2007) believed, time management includes designing and planning of the study time and Terry (2002) and Vecchio, (2003) say that these individual can do the tasks more effectively and consistently.

These results are also supported in the studies such as Chen (2002); Filcher and Miller (2000). The significance of the simple relationship between help seeking and academic achievement has also been in line with the studies such as Salmons (2008), which has revealed that co-operation learning leads to improvement in learning and enhancement of academic achievement.

Salmons (2008) argues that help-seeking is also another efficient way to overcome difficulties, and allows for decreasing psychological depressions and enhancement of learning and problem solving skills and it is an effective learning strategy that provides the basis for successful learning when students lack the adequate knowledge or understanding required for solving the problem by themselves. This characteristic not only helps learners to satisfy their learning needs immediately, but also is an effective way for elevating and improvement of their performances. Results from the study of Life (2011) also indicate that, in learning English, Japanese students prefer teamwork. However, studies show that there are some learners who possess the ability and skill of help-seeking while others lack such a skill (Wilson & Deane, 2002). By considering the role of problem



solving and team work on the academic achievement, can be said that the findings of previous studies supported the results of present research.

All in all, in the line of purposes of current study following results were obtained. Then pedagogical implications are suggested:

- Resource management strategies subscales were significant predictor for academic achievement of English language students. By considering resource management strategies component could be learned and developed, it is suggested that through workshops, seminars, and other ways; resource management strategies would be trained to the students.
- Among subscales of resource management strategies, effort regulation has highest share in predicting academic achievement. So with regarding the construction of effort regulation students should apply following recommends: They work hard to do well in their class even if don't like what they are doing; when course work is difficult, they neither give up nor only study the easy parts; even when course materials are dull and uninteresting, they manage to keep working until they finish.

Interpretations of the findings of this study also lead the researcher to several recommendations for future and further research: Future study in this area would better carry out in the experimental method. Academic achievement would be premeasured at the beginning of term, and then components of motivated strategies for learning would be trained during term to the students, at the end academic achievement would be post measured. In this method, the effect of motivated strategies for learning on academic achievement will specify well.

## References

- Abbasi, M., Ahmadi, G., & Lotfi, A. (2009). Problems of teaching and learning of English language among high school students from teachers' perspectives. *Knowledge & Research in Applied Psychology*, 23, 141–156.
- Arnold, E., & Pulich, M. (2004). Improving productivity through more effective time management. *The Health Care Manager*, 23(1), 65–70.
- Bigdeli, M. (2010). *Factors affecting on drop in English language learning in high school*. Ardabil: Ministry of Education.
- Britton, B. K., & Tessor, A. (1991). Effects of time management practices on college grades. *Journal of Educational Psychology*, 83(3), 405–410.



- Campbell, M. M. (2007). Motivational systems theory and the academic performance of college students. *Journal of College Teaching & Learning*, 4(7), 78–85.
- Canagarajah, A. S. (2002). *The geopolitics of academic writing*. Pittsburgh, PA: University of Pittsburgh.
- Chen, C. S. (2002). Self-regulated learning strategies and achievement in an introduction to information system course. *Information Technology, Learning, and Performance Journal*, 20(1), 11–24.
- Filcher, C., & Miller, G. (2000). Learning strategies for distance education learners. *Journal of Agricultural Education*, 41(1), 41–53.
- Gran-Moravec, M. B., & Hughes, C. M. (2005). Nursing time allocation and other considerations for staffing. *Nursing Health and Sciences*, 7, 126–133.
- Karabenick, S. A., & Knapp, J. R. (1991). Relationship of academic help seeking to the use of learning strategies and other instrumental achievement behavior in college students. *Journal of Educational Psychology*, 83, 221–130.
- Kivinen, K. (2003). *Assessing motivation and the use of learning strategies by secondary school students in three international schools*. Unpublished Dissertation, University of Tampere, Finland. Retrieved June 15, 2011, from <http://acta.uta.fi/pdf/951-44-5556-8.pdf>.
- Kosnin, A. M. (2007). Self-regulated learning and academic achievement in Malaysian undergraduates. *International Education Journal*, 8(1), 221–228.
- Life, J. (2011). Motivation and EFL University students in North-East Asia. *The Asian EFL Journal Quarterly*, 13(3), 11–41.
- Newman, R. S., & Schwager, M. T. (1993). Student perceptions of the teacher and classmates in relation to reported help seeking in math class. *Elementary School Journal*, 94, 3–17.
- Pakdaman, M., Farzad, V., Sarmad, Z., & Khanzadeh, A. (2007). The role of help-seeking on mathematics problem solving. *Educational Quarterly Journal*, 95, 119–137.
- Pintrich, P. R. (2003). A motivational science perspective on the role of student motivation in learning and teaching contexts. *Journal of Educational Psychology*, 95, 667–686.
- Pintrich, P. R. (2004). A conceptual framework for assessing motivation and self-regulated learning in college students. *Educational Psychology Review*, 16, 385–407.
- Rickwood, D., Deane, F. D., Wilsons, C. J., & Ciarrochi, J. (2005). Young people's help-seeking for mental health problems. *Australian e-Journal for the Advancement of Mental Health*, 4(3), 512–517.
- Sadighi, F., & Zarafshan, M. (2006). Effects of attitude and motivation on the use of language learning strategies by Iranian EFL university students. *Journal of Social Sciences & Humanities of Shiraz University*, 23(1), 2–19.
- Salmons, J. E. (2008). Taxonomy of collaborative e-learning. In L. E. Tomei, (Ed.), *Encyclopedia of international technology curriculum integration* (pp. 839–846), Information Science Reference.



- Schunk, D. H., & Zirnnerman, B. J. (1994). *Self-regulation of learning and performance: Issues and educational applications*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Sheikholeslami, R., & Khayer, M. (2006). The relation between motivational beliefs and English language learning strategies among university students. *Journal of Psychology*, 10(1), 22–33.
- Taplin, M., Yum, J. C. K., Jegede, O., Fan, R. Y. K., & Chan, M. S. -Ch. (1999). *Help-seeking strategies used by high-achieving and low-achieving distance education students*. In Asian Association of Open Universities, October 14th–17th, (pp. 153–159), Beijing.
- Tardy, C. (2004). The role of English in scientific communication: Lingua franca or Tyrannosaurusrex? *Journal of English for Academic Purposes*, 3, 247–269.
- Terry, P. S. (2002). *The effect of online time management practice on self-regulated learning and academic self-efficacy*. Unpublished dissertation of curriculum and instruction, Virginia University.
- Vecchio, R. (2003). *Organizational behavior: Core concepts* (5th ed.). Mason OH: South-Western.
- Williams, M., & Burden, R. L. (2000). *Psychology for language teachers* (3rd ed.). Cambridge: Cambridge University press.
- Zimmerman, B. J., & Martinez-Pons, M. (1986). Development of a structured interview for assessing student use of self-regulated learning strategies. *American Educational Research Journal*, 23, 614–628.

### Bio Data

**Maryam Mohammadi** holds MA degree of teaching English course from the department of English, Islamic Azad University, Ardabil, Iran. She got her BA degree in English Language Translation from Payam-e-Noor University of Ardebil, Iran.

**Afsar Rouhi** is assistant professor of TEFL at Payam-e-Noor University, Iran. He got his Ph.D. in TEFL from University of Tehran, Iran in 2006. Since then he has been intimately involved in teaching SLA courses at Payam-e-Noor University and supervising MA and PhD theses and dissertations.

**Mehran Davaribina** is the faculty member in the Department of English at Islamic Azad University in Ardabil, Iran.



## **Evidence of Sociocultural Transfer in Iranian Culture: A Case of Suggestion Act**

**Maryam Sharafadini<sup>1</sup>**

*Ferdowsi University of Mashhad, Iran*

### **Abstract**

This study is an investigation into sociocultural transfer employed by Iranian learners of English as a foreign language while making suggestion acts. To this end, 150 Persian natives and 150 Iranian English learners participated in this study and a Discourse Completion Task (DCT) of six situations was applied to extract learners' suggestion responses. The Iranian EFL learners' suggestion performances were compared to those of Persian natives to detect whether sociocultural language transfer has occurred. In fact, this study investigates similarities and differences between English and Iranian culture. The study findings indicated evidence of sociocultural transfer especially in regard to cultural respects.

**Keywords:** Culture, Pragmatic competence, Sociocultural transfer, Speech act, Suggestion speech act

### **1. Introduction**

Culture and language are closely interconnected in a way that sociocultural conventions designate our way of thinking and speaking (Sapir, 1949; Whorf, 1956, as cited in Liu, 1995). Different nations in the world possess their own particular cultures which demonstrate some "universalities and particularities" (Wei, 2009. p. 1). Many people experience communication breakdown or even communication conflict in their cross-cultural interactions with people from different language backgrounds. Speech performances are organized by particular cultural and social constraints (Gumperz & Hymes, 1972, as cited in Al-Issa, 2003). Therefore, people from various sociocultural backgrounds employ their own sociocultural values in their interactions and deviate from L2 culturally different forms. According to Thomas (1983, as cited in Wu, 2007), sociopragmatic failure occurs when non-native speakers fail to opt for appropriate codes due to their lack of recognizing cultural differences. Therefore, interlocutors should attend to cultural mode of the message in addition to its form. In other words, a successful communication demands grammatical as well as

---

<sup>1</sup> Email: maryam.sharafadini@gmail.com



sociolinguistic competence (Canale & Swain, 1980; Richards, 1980). Speech act studies will yield beneficial information about sociocultural values underlying societies' communicative system.

Interlanguage Pragmatics (ILP) is one of the significant concepts in second language acquisition in which learners' pragmatic competence is investigated (Kasper, 1992, as cited in Wannaruk, 2008). According to Kwon (2003), EFL learners' speech act samples vary from that of native speakers due to their incomplete knowledge of L2 sociolinguistic norms. Many pragmatic studies (e.g. Allami & Naeimi, 2010; Blum-Kulka, 1982) revealed that sociocultural values differ from one culture to another and approved the existence of pragmatic transfer (Kasper, 1992; Keshavarz, Eslami-Rasekh & Ghahraman, 2006). Sociolinguists relates such miscommunications to different "value systems" evoked by interlocutors' L1 cultural conventions (Chick, 1996, p. 329, as cited in Urano, 2000). Different value systems are realized in different speech act interpretations and productions. Pragmatic transfer has been investigated in regard to different speech acts in various cultures (Byon, 2004; Hassall, 2003; Jaworski, 1994; Jalilifar, Hashemian & Tabatabaee, 2011), whereas suggestion speech act has absorbed scant attention.

Suggestion speech act is frequently utilized in our daily interactions. It is regarded as a face-threatening act (Brown & Levinson, 1987), and belongs to directive type of speech acts in which we get people to do something. In a suggestion, a possibility is represented by a speaker for the hearer (Sum-Hung Li, 2010). Suggestion speech act realizations vary in different cultures; direct and indirect suggestion realizations may not be differentiated by EFL learners. Therefore, it seems essential to investigate suggestion act within various cultures. This study endeavors to investigate how Iranian EFL learners' suggestion realizations are transferred from their native culture while interacting in English. As a matter of fact, a comparative analysis is conducted in regard to sociocultural transfer. Therefore, the current study intends to explore Iranian learners' sociocultural transfer in regards of cultural deep levels. Furthermore, several extracts are proposed to reveal the culturally specific variations within two cultures.

### *1.1. Theoretical framework*

Austin (1962) introduced three levels of locutionary, illocutionary, and perlocutionary for each speech act utterance. In fact, locutionary act indicates the literal meaning of an utterance; illocutionary act emphasizes the speaker's intention and perlocutionary act deals with the effect exerted by stating an utterance. On the other hand, Searle (1976) classified pragmatic speech acts into five categories of representative, directive, commissive, expressive, and declaration which suggestion act belongs to directives. A suggestion is a directive type of speech act in which the speaker offers an idea or action for other people to consider or suggests an opinion about what



other people should do in a certain situation and believe the mentioned action is in the benefit of the hearer (Jiang, 2006).

Speech act universality is one of the major issues in speech act studies. Are speech acts universal and if so, can we consider the same speech act taxonomy for all languages? Many studies demonstrated non-native learners' gap in their pragmatic performance in comparison with native speakers (e.g. Alcon Soler & Codina Espurz, 2002; Bardovi-Harlig, 2001; Wannaruk, 2008) which may be due to their incompetence in L2 sociolinguistic rules (Kwon, 2003, as cited in Wannaruk, 2008). Consequently, they compensate this gap utilizing their native sociolinguistic norms while communicating in L2; hence, pragmatic transfer may occur. "Pragmatic transfer in interlanguage pragmatics shall refer to the influence exerted by learners' pragmatic knowledge of languages and cultures other than L2 on their comprehension, production, and learning of L2 pragmatic information" (Kasper, 1992, p. 209, as cited in Bou Franch, 1998). Learners' insufficient knowledge of L2 sociolinguistic rules and their pragmatic transfer from their first language may lead to communication breakdown. As Richards (1980, p. 430) stated

Transfer of features of first language conversational competence into English may have much more serious consequences than errors at the levels of syntax or pronunciation, because conversational competence is closely related to the presentation of self; that is, communicating an image of oneself to others. Therefore, transferring from their L1, learners might be regarded as rude or even inconsiderate which indicates the necessity of implementing instruction in pragmatics for the non-native learners to acquire the pragmatic competence.

### *1.2. Empirical framework*

As Liu (1995) stated most scholars have investigated the sociocultural transfer on surface structures, whereas deep sociocultural values did not receive much attention. Liu (1995) examined deep sociocultural transfer within Chinese students while responding to compliments in English. In fact, the study was conducted in two phases: first, compliment responses of 20 students were observed. Second, a questionnaire consisting of 8 situations was distributed among the same students after observation. The findings revealed that deep sociocultural transfer was evident in the speech of even proficient Chinese students. Furthermore, the need of L2 learners to develop sociocultural competence was discussed. In another study, Al-Issa (2003) not only explored concept of sociocultural transfer but also its motivating factors were investigated in regard to Jordanian EFL refusal patterns. The research data was collected by a DCT based on observational field note data. Moreover, administrating the DCT, follow up interviews were conducted. EFL learners' refusal



utterances were compared with English native speakers and native speakers of Arabic. The findings demonstrated sociocultural transfer in several respects such as choice of selecting semantic formulas, length of responses, and content of semantic formulas. Additionally, several motivating factors of sociocultural transfer were discussed. On the other hand, Fernández Amaya (2008) focused on the pragmatic failure and proposed several examples to investigate how pragmatic failure can impede a communication thoroughly. He considered the need to incorporate the pragmatic and cultural issues of L2 learning into pedagogical tools and offered several techniques to enhance the pragmatic competence. Although several studies have been conducted on suggestion act (Banerjee & Carrell, 1988; Sum-hung Li, 2010), to date no study has intended to discover the sociocultural transfer of suggestion speech act.

### *1.3. Purpose of the study*

According to Liu (1995, p. 253), most scholars investigated L2 learners' sociocultural transfer in regard to "surface-structure transfer" and have neglected "deep-structure transfer", in which the learners transfer their home sociocultural values into their target languages. Deep sociocultural transfer not only exists, but also causes serious consequences. Learners committing such transfers might be perceived rude or disrespectful in their I2 interactions with native speakers (Wannurak, 2008). Therefore, this study intends to explore how Iranian EFL learners' suggestion speech act is transferred from their native culture while they are interacting in English. Furthermore, the following study question is probed in this study.

**Q1:** What are the cultural and pragmatic similarities and differences on the deep levels between Iranian EFL learners and Persian natives in the production of suggestion speech act?

## **2. Method**

### *2.1. Participants*

The study participants involved two groups. First, they involved 150 Iranians who were learning English in language institutes in Mashhad- a city in Iran. They were 75 males and 75 females aged 17 to 50 with different socio-economical backgrounds. The participants consisted of 50 intermediate, 50 upper intermediate, and 50 advanced learners. Their levels of proficiency were determined by the language institutes in which they were studying. They had not spent a long time in an English native country and their only access to English was through the English classes. In fact, English in Iran has no out-of-class application for communication. Therefore, it is considered as a foreign language in Iran.



Second group consisted of a total of 150 Iranian native university students who were studying different majors such as mechanical engineering, civil engineering, statistics, politics, history, social sciences, etc. in Ferdowsi University of Mashhad. They involved 75 males and 75 females aged from 17 to 45 distributed homogenously with different socio-economical backgrounds. As a matter of fact, since studying a second language may influence learners' L1, language students of Persian, English, Arabic, French, and Russian were excluded. Moreover, none of them has been exposed to another language for a long period of time.

## *2.2. Instrumentation*

To glean the research data, an English DCT and its related Persian version were applied in which six natural situations were defined elaborately. According to Cohen (1996), one of the straightforward means to glean the pragmatic data is DCT and if it is well-prepared, it reveals how respondents activate their pragmatic knowledge (Martinez-Flor, 2006). Participants were required to respond as they would say in their daily conversations (see Appendix I and II). The DCT was designed based on the guidelines laid down by Banerjee and Carrell (1988), Martinez-Flor (2005), Martinez-Flor (2006), Martinez-Flor and Alcon Soler (2007), and Martinez-Flor and Fukuya (2005). As Hudson, Detmer and Brown (1995, as cited in Martinez-flor & Alcon soler, 2007) stated the content of the DCT should be familiar to the students in terms of context. Thus, three DCT situations were involved with familiar daily interactions, whereas the other situations were related to students' educational affairs because all the participants were students. Moreover, according to Lorenzo-Dus (2001), cultural familiarity to the situations is a crucial element in the quality of responses; thus, "cultural unfamiliarity" (Lorenzo-Dus, 2001, p. 112) was eliminated as much as possible. Since the DCT questionnaires did not have the sufficient interaction available in authentic discourse, a short dialogue was added to each situation. Regarding the English DCT, following Matsumura's (2001) suggestion, learners were asked to imagine themselves in a foreign country studying English language. According to the guidelines provided by Martinez-Flor and Alcon-Soler (2007), the test instructions were given in L1 since a thorough understanding of how task should be performed is necessary for the learners. Furthermore, including both male and female participants, the situations were devised in a gender neutral way and regarding the status, to have a representative sample of authentic discourse, the situations involved inferior, equal and superior relationships.

The efficacy of questionnaires was examined in a pilot study which was administrated in two phases: in the first phase, 20 EFL learners and in the second phase 20 Iranian university students took part among which some of the participants were interviewed. Each interview lasted about 10-15 minutes and the interviewees' beneficial points were recorded for further investigations. A team



of specialists in L2 were asked to substantiate the content validity of the both DCTs which were modified accordingly and their ambiguities and obstacles were eliminated.

### 2.3. Procedure

The English DCT was responded by Iranian EFL participants studying English at different language institutes of Mashhad. They completed the questionnaire in the last 15 minutes of their class with the permission of their teachers. The necessary instruction was presented by the researchers in English and one example was given in their first language. The Persian native participants who were university students of Ferdowsi University responded the Persian DCT individually which took them about 15 minutes. They were first asked for their permission and then were provided with the necessary instruction. Since the task required to be understood thoroughly, an example was given by the researchers. After gleaning the research data, responses were interpreted qualitatively on deep cultural and pragmatic levels.

## 3. Results

Regarding sociocultural transfer, there exist several cultural pieces of evidence to prove such transfer including “define relationship” (Al-Issa, 2003, p. 585), friendship, ritual politeness (*'tæ'arof'*), excuse me, and cultural indirectness. Iranian EFL participants employed “define relationship” expressions in their suggestions in which the speaker calls the other interlocutor with his title. In the below situation, both groups started their suggestion acts with first defining the relationship for instance:

Persian EFL learner: My dear neighbor, I suggest you to see the expiration date of those potato chips.

Persian natives: *Hamsayeye aziz mitoonam chipsatoono bebinam?/ Dear neighbor, may I have a look at your chips?*

The next issue to be considered is the concept of friendship. Iranian English learners, communicating in their L2, have transferred their L1 friendship values to their L2 communication, for instance:

Persian EFL learner: I think you need a little rest; I can help you in searching and you can rest a little.



Persian Natives: *Boro esterahart kon choon talash o jostojoo dar in zaman fayedeye kafi ro nakhahd dasht. Man khodam karhayat ra anjam midam.* / Go and rest because trying to search now is not fruitful. I'll do instead of you.

Closely related to this concept, is the concept of '*tæ'arof*' or ritual politeness. Etymologically, Koutlaki (2002, p. 1740) defined '*tæ'arof*' as "mutual recognition". Since there is no direct equivalent for '*tæ'arof*' in English, Koutlaki (2002) stated different researchers who proposed alternative equivalents for this term such as "ritual courtesy" (Beeman, 1986, p. 56), and "polite verbal wrestling" (Rafiee, 1992, p. 96). It has got a very complex structure and plays a crucial role in the discussion of Persian politeness system. The sign of '*tæ'arof*' can be observed in Iranian EFL cultural transfer, for instance:

Persian EFL learner: Be my guest, I'll go and buy a cup of coffee for you.

Persian natives: *Ostad, age khaste bashid mayeye maserat kahad bood ke noskheyen arzantar ro baraye shoma tahiye konam.* / Professor, if you want, it would be a great pleasure to prepare the cheaper version of the book for you.

In the Iranian culture, Iranians utilized '*bebakhshid*' (excuse me) as a softening device to maintain the interlocutors' face. In other words they apologize, though they have not committed any mistake. '*Bebakhshid*' is the most commonly utilized apology strategy within Persian-speaking people (Chamani & Zareipur, 2010). EFL learners translated this concept literally into "excuse me" which was employed in their suggestions quite frequently. For instance:

Persian EFL learner: Excuse me neighbor, these potato chips are expired and they can be harmful to eat.

Persian natives: *Bebakhshid hamasayeh tarikhe enghezaye in chipsaha gozashte ast.* / Excuse me my neighbor, the expiration date of these chips has passed.

The other point to be looked upon in regard to sociocultural transfer is the indirectness. Iran's culture is viewed as an indirect culture. Iranian EFL learners transferred such indirectness to L2 interactions in which in contrary to Eastern cultures the codes are interacted explicitly and a direct style of communication is preferred. For instance:

Persian EFL learner: I give her one of my candies and she will understand what a terrible taste it has. (Situation 6 in which a little girl should be suggested about a terribly tasted chocolate.)



Persian natives: *Salam hamsayeh. Man ham asheghe chipsam ama be sharti ke khoshmaze va salem bashe. Be manam ye chips bedin. Bad sari be tarikhe oon negah mikonam va be forooshande pas midam. Hatman hamsaye ham negah mikone va mifahmeh.* Hi my neighbor, I love chips too only if they are delicious and healthy. Then immediately, I look at its expiration date and return it to the shopkeeper. Surely, my neighbor is watching and will understand. (Situation 4 in which participants should suggest their neighbor about expired chips.)

Persian culture is closely tied with the concept of religion. As Al-Issa (2003) mentioned Muslims regard Islam as one of the integral parts of their everyday life; therefore, Persian natives also demonstrated their faith in Allah in making suggestion strategies, for instance:

Persian natives: *enshallah ke zoodtar tamoom mishe. Ye kam esterahat kon, bad edameh bedeh.* Godwilling, it finishes very soon. Rest a little while and then continue.

It is worth mentioning that none of Persian EFL learners have demonstrated such cultural norm while making suggestions in their L2.

#### 4. Discussion

As it was mentioned before, several pieces of cultural evidence were proposed to prove the existence of sociocultural transfer including “define relationship” (Al-Issa, 2003, p. 585), friendship, ritual politeness (*'tæ'arof*), excuse me, and cultural indirectness. Regarding “define relationship”, Hofstede (1980, as cited in Tavakoli, Keenan, & Crnjak-Karanovic, 2003) proposed four dimensions to classify all cultures of world including power distance, uncertainty avoidance, individualism, and masculinity among which power distance and individualism provide beneficial insights into our discussion in terms of illuminating Iranian cultural norms. Power distance, is “the extent to which the less powerful members of institutions and organizations within a country expect and accept that power is distributed unequally” (Hofstede, 1997, p. 28, as cited in Tavakoli et al., 2003). Such hierarchy associates with Iranian sociocultural values, whereas the Westerns prefer the equality. As Shang-chao (2008) indicated, the Western society emphasizes egalitarianism and assertiveness; on the contrary, Easterns involve a non-egalitarian society who value hierarchical structure. In fact, according to Eslami-rasekh, Tavakoli, and Abdolrezapour (2010), Iranian culture is regarded as hierachal on the basis of Scollon and Scollon's (2001) category, in which social order should be preserved, whereas American culture employs a deference politeness system in which the interlocutors are given equal social level and equal rights. Therefore, Iranians' society has a hierachal culture wherein they attempt to maintain such hierarchy using various politeness strategies in order to show respect. One of the means to reveal such respect is via “defining



relationships" in which the speaker calls the other interlocutor with his title. As it is obvious a negative transfer has occurred because this formula is absent in the English native sample. Therefore, our findings regarding "defining relationship" are in line with those of Al-Issa (2003) who reported defining relationship strategy comparing Arab English learners with English natives.

Regarding friendship, Iranians share a collectivistic culture (Chang, 2008); thus, group harmony is regarded as an essential constituents in such cultures, whereas the westerns belong to an individualistic culture in which "the ties between individuals are loose" (Hofstede, 1997, p. 51, as cited in Tavakoli et al., 2003). The concept of friendship as Al-Issa (2003) stated is an inseparable part of social structures for instance, in such Eastern societies, friendship norms demand to fulfill certain obligations such as offering help and doing whatever is possible to help a friend (Al-Issa, 2003). Therefore, our findings are in line with Byon's (2004) findings who described his Korean participants as collectivistic as well.

Regarding '*tæ'arof*' which is defined as "a tool for negotiating interactants' relationships" (Koutlaki, 2002, p. 1740). As a matter of fact, Iranians offer '*tæ'arof*' to value the other interlocutors' face in order to show respect. In fact, in '*tæ'arof*' a favor is taking place and the interlocutors recognize that it is not polite to accept this favor. They do not sincerely mean to do the favor but demonstrate '*ehteram*' (respect). Therefore, as Eslami-Rasekh (2005, as cited in Keshavarz et al., 2006) stated in Iranian culture, being regarded polite is more important than being sincere.

Regarding use of "bebakhshid" in the Iranian culture, as Eslami-Rasekh (2005, as cited in Keshavarz et al., 2006) stated, being polite is so significant that might be posited in a higher rank than being sincere. As a matter of fact, concept of face is so crucial that exerting face-threatening speech acts such as suggestions, Iranians utilized '*bebakhshid*' (excuse me) as a softening device to maintain the interlocutors' face and demonstrate the necessary respect for them. Moreover, Iranian EFL learners translated the word '*bebakhshid*' literally to "*excuse me*" and applied that to mitigate the enforcement of suggestion acts. Therefore, our findings are in line with those of Eslami-Rasekh (2004), Eisenstein and Bodman (1986) which literal word-for-word translation in their participants' speech samples was found.

Regarding indirectness, according to Allami and Naeimi (2010), Iranian culture is regarded as a low-context culture (Hall, 1976) in which communication is coded implicitly. This is the reason why Iran's culture is viewed as an indirect culture. Iranian EFL learners transferred such indirectness to L2 interactions in which in contrary to Eastern cultures demand a high context culture (Hall, 1976); the codes are interacted explicitly and a direct style of communication is preferred.



Regarding Religion, Hofstede and Bond (1988, as cited in Tavakoli et al., 2003) introduced another category named long term orientation in contrary to short term orientation. Long term cultures regard future rewards and thrift, whereas short term cultures take both past and present into consideration; face perseverance, social conventions and religion are extremely appreciated in such cultures. This can be demonstrated in Iranians' strong ties with religion and traditional conventions (e.g. '*ehteram*', friendship, '*tæ'arof*'). Generally our study is in line with the bulk of studies (Al-Issa 2003; Ghawi, 1993; Liu, 1995; Yu, 2011) in which instances of sociocultural transfer were detected in regard to different cultures.

This study entails some pedagogical implications. The present study findings demonstrate that speech acts studies can bring about fruitful sociocultural information in cross-cultural interactions. The main contribution of such information lies in comprehending aspects of communicative competence. In fact, the study findings emphasize the need for L2 learners to acquire not only the grammar, but also the sociocultural rules to be able to communicate competently in L2 community. From a cross-cultural point of view, they also indicate those areas in which communication breakdown may occur. If both groups bear different sociocultural expectations in mind, an efficient communication between Iranian and English natives would not occur. For instance, we can consider the case of "excuse me". Iranians utilize this softener to respect the speakers' face, whereas the English natives might regard it as an insincere and strange kind of act. Thus, this fact is quite clear that teachers and learners should move from surface patterns of communication to in-depth sociocultural values.

This research suffers from some limitations. Therefore, some suggestions are made for future studies. This research has been conducted on a small scale and the study findings cannot be generalized to all Iranians. Thus, further studies need to be conducted with a larger different sample. In addition to DCTs, ethnographic methods also could be applied to offer more in-depth data regarding Iranian sociocultural values. Furthermore, several variables such as age, social class, and educational background may be attended. Finally, further studies can be conducted on other cultures as well to demonstrate a cross-cultural picture of sociocultural values.

## References

- Alcon Soler, E., & Codina Espurz, V. (2002). Practice opportunities and pragmatic change in a second language context: The case of requests. *ELIA*, 3, 123-138.
- Al-Issa, A. (2003). Sociocultural transfer in L2 speech behaviors: Evidence and motivating factors. *International Journal of Intercultural Relations*, 27(5), 581- 601.



- Allami, H., & Naeimi, A. (2010). A cross-linguistic study of refusals: An analysis of pragmatic competence development in Iranian EFL learners. *Journal of Pragmatics*, 43(1), 385-406.
- Austin, J. (1962). *How to do things with words*. Oxford: Oxford University Press.
- Banerjee, J., & Carrell, P. L. (1988). Tuck in your shirt, you squid: Suggestions in ESL. *Language Learning*, 38(3), 313-364.
- Bardovi-Harlig, K. (2001). Evaluating the empirical evidence: Grounds for instruction in pragmatics. In K. R. Rose, & G. Kasper (Eds.), *Pragmatics in language teaching* (pp. 13-32). Cambridge: Cambridge University Press.
- Beeman, W. O. (1986). *Language, status and power in Iran*. Indiana University Press.
- Blum-Kulka, S. (1982). Learning how to say what you mean in a second language: A study of speech act performance of learners of Hebrew as a second language. *Applied Linguistics*, 3(1), 29-59.
- Brown, P., & Levinson, S. (1987). *Politeness: Some universals in language usage*. Cambridge: Cambridge University Press.
- Bou Franch, P. (1998). On pragmatic transfer. *Studies in English Language and Linguistics*, 0, 5-20.
- Byon, A. S. (2004). Sociopragmatic analysis of Korean requests: Pedagogical settings. *Journal of Pragmatics*, 36(9), 1673-1704.
- Canale, M., & Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics*, 1(1), 1-47.
- Chamani, F., & Zareipur, P. (2010). A cross-cultural study of apologies in British English and Persian. *Studies in Linguistics*, 36(1), 133-153.
- Chang, Y. (2008). How to say no: An analysis of cross-cultural difference and pragmatic transfer. *Language Sciences*, 31(4), 477-493.
- Chick, J. K. (1996). Intercultural communication. In S. L. McKay, & N. H. Hornberger (Eds.), *Sociolinguistics and language teaching* (pp. 329-348). Cambridge: Cambridge University Press.
- Cohen, A. D. (1996). Speech acts. In S. L. McKay, & H. N. Hornberger (Eds.), *Sociolinguistics and language teaching* (pp. 383-420). Cambridge: Cambridge University Press.
- Eisenstein, M., & Bodman, J. W. (1986). "I very appreciate": Expressions of gratitude by native and nonnative speakers of American English. *Applied Linguistics*, 7(2), 167-185.
- Eslami-Rasekh, Z. (2004). Face-keeping strategies in reaction to complaints: English and Persian. *Journal of Asian Pacific Communication*, 14(1), 179-195.
- Eslami-Rasekh, Z. (2005). Invitations in Persian and English: Ostensible or genuine? *Intercultural Pragmatics*, 2(4), 453-480.
- Eslami-Rasekh, A., Tavakoli, M., & Abdolrezapour, P. (2010). Certainty and conventional indirectness in Persian and American request forms. *Medwell Journal, the Social Sciences*, 5(4), 332-339.



- Fernandez Amaya, L. (2008). Teaching culture: Is it possible to avoid pragmatic failure? *Revista Alicantina de Estudios Ingleses*, 21, 11-24.
- Ghawi, M. (1993). Pragmatic transfer in Arabic learners of English. *El Two Talk*, 1(1), 39-52.
- Gumperz, J., & Hymes, D. (Eds.). (1972). *Directions in sociolinguistics: The ethnography of communication*. New York: Holt, Rinehart and Winston.
- Hall, E.T. (1976). *Beyond Culture*. New York: Doubleday.
- Hassall, T. (2003). Requests by Australian learners of Indonesian. *Journal of Pragmatics*, 35, 1903-1928.
- Hofstede, G. (1997). *Cultures and organizations: Software of the mind*. London: McGraw-Hill.
- Hofstede, G. (1980). *Culture's consequences: International differences in work-related values*. Beverly Hills, CA: Sage Publications.
- Hofstede, G., & Bond, M. H. (1988). The Confucius connection: From cultural roots to economic growth. *Organizational Dynamics*, 16(4), 4-21.
- Hudson, T., Detmer, E., & Brown, J. D. (1995). *Developing prototypic measures of cross-cultural pragmatics*. Honolulu: University of Hawaii Press.
- Jalilifar, A., Hashemian, M., & Tabatabaei, M. (2011). A cross-sectional study of Iranian EFL learners' request linguistic devices. *Journal of Language Teaching and Research*, 2(4), 790-803.
- Jaworski, A. (1994) Pragmatic failure in a second language: Greeting responses in English by Polish students. *IRAL*, 32(1), 41-55.
- Jiang, X. (2006). Suggestions: What should ESL students know? *Elsevier*, 34(1), 36-54.
- Kasper, G. (1992). Pragmatic transfer. *Studies in Second Language Acquisition*, 8(3), 203-231.
- Keshavarz, M. H., Eslami-Rasekh, Z., & Ghahraman, V. (2006). Pragmatic transfer and Iranian EFL refusals: A cross-cultural perspective of Persian and English. In K. Bardovi-Harlig, J. C. Felix-Brasdefer, & A. S. Omar (Eds.), *Pragmatics and language learning* (pp. 359-402). University of Hawai'i Press: Honolulu.
- Koutlaki, S. (2002). Offers and expressions of thanks as face enhancing acts: Tæ'arof in Persian. *Journal of Pragmatics*, 34(12), 1733-1756.
- Kwon, J. (2003). *Pragmatic transfer and proficiency in refusals of Korean EFL learners*. (Doctoral dissertation). Boston University, United States.
- Liu, D. (1995). Sociocultural transfer and its effect on second language speakers' communication. *International Journal of Intercultural Relations*, 19(2), 253-265.
- Lorenzo-Dus, N. (2001). Compliment responses among British and Spanish university students: A contrastive study. *Journal of Pragmatics*, 33(1), 107-127.
- Martinez-Flor, A. (2005). A theoretical review of the speech act of suggesting: Towards taxonomy for its use in FLT. *Revista Alicantina de Estudios Ingleses*, 18, 167-187.



- Martinez-Flor, A. (2006). Task effects on EFL learners' production of suggestions: A focus on elicited phone messages and Emails. *Journal of English and American Studies*, 33, 47-64.
- Martinez-Flor, A., & Alcon Soler, E. (2007). Developing pragmatic awareness of suggestions in the EFL classroom. *Canadian Journal of Applied Linguistics*, 10(1), 47-76.
- Martinez-Flor, A., & Fukuya, Y. J. (2005). The effects of instruction on learners' production of appropriate and accurate suggestions. *System*, 33(3), 463-480.
- Matsumura, S. (2001). Learning the rules for offering advice: A quantitative approach to second language socialization. *Language Learning*, 51(4), 635-679.
- Rafiee, A. (1992). Variables of communicative incompetence in the performance of Iranian learners of English and English learners of Persian. Doctoral dissertation, University of London.
- Richards, J. C. (1980). Conversation. *TESOL Quarterly*, 14(4), 413-432.
- Sapir, E. (1949). *Selected writings in language, culture and personality*. Berkeley and Los Angeles: University of California Press.
- Scollon, R., & Scollon, S. (2001). *Intercultural communication*. Oxford: Blackwell Publishers Ltd.
- Searle, J. (1976). A classification of illocutionary acts. *Language in Society*, 5(1), 1-23.
- Shang-chao, M. (2008). Study on the differences of speech act of criticism in Chinese and English. *US-China Foreign Language*, 6(3), 74-77.
- Sum-hung Li, E. (2010). Making suggestions: A contrastive study of young Hong Kong and Australian students. *Journal of Pragmatics*, 42(3), 598-616.
- Tavakoli, A. A., Keenan, J. P., & Crnjak-Karanovic, B. (2003). Culture and whistle blowing an empirical study of Croatian and United States managers utilizing Hofstede's cultural dimensions. *Journal of Business Ethics*, 43, 49-64.
- Thomas, J. (1983). Cross-cultural pragmatic failure. *Applied Linguistics*, 4(2), 91-112.
- Urano, K. (2000). Negative pragmatic transfer in compliment responses by Japanese learners of English: A research proposal. *Shinshu University Research on Communication and Language Education (SURCLE)*, 2, 27-38.
- Wannaruk, A. (2008). Pragmatic transfer in Thai EFL refusals. *RELC*, 39(3), 318-337.
- Wei, X. (2009). On negative cultural transfer in communication between Chinese and Americans. *Journal of Intercultural Communication*, 21. Retrieved August, 12, 2011, from <http://www.immi.se/jicc/index.php/jicc/article/view/48>.
- Whorf, B. L. (1956). *Language, thought, and reality: Selected writings*. Cambridge, Mass: MIT Press
- Yu, M. (2011). Learning how to read situations and know what is the right thing to say or do in an L2: A study of socio-cultural competence and language transfer. *Journal of Pragmatics*, 43(4), 1127-1147.



### Bio Data

**Maryam Sharafadini** obtained her M.A degree in teaching English from Ferdowsi University of Mashhad. She has been teaching English as an ESP course in Shahid Sadoughi University of Yazd as well as in different language institutes. Her research interests are mainly in sociolinguistics and psycholinguistics.



## A New Model for General English Textbooks' Design: Inviting Philosophy to the Scene

Abbas Moradan  
Mohammad Amerian<sup>1</sup>

*Semnan University*

### Abstract

This paper tries to introduce and elaborate on a new content model for the General English textbooks' design. In general, it is seen that in syllabus design and material preparation, some challenging topics have been ignored altogether because it was thought that they will encompass deviations from the educational syllabus or even endanger the students' ideologies or beliefs. Philosophical and epistemological issues are among such kinds of subjects. Although existing all over the world, this is mainly seen in the countries in which religious and cultural traditions are mostly valued and emphasized. The proposed model in this article is based on the actual observational needs-analysis by the researchers and has taken different fields of study (as History, Philosophy, Cosmology and Epistemology) into account through a gradual procedure. The model asserts that if regulated well; challenging topics will not deviate the learners, but in fact, stimulate their wisdom and flourish their judgment and outlook. In brief, they will open new horizons for the intellectual development of learners while they are engaged with language learning.

**Keywords:** Syllabus design, Material preparation, General English textbooks, Challenging subjects, Philosophical topics

### 1. Introduction

One of the key issues in language learning and teaching is syllabus design and curriculum development. There exist a number of different definitions for syllabus design. Graves (1996) defines syllabus as the specification and ordering of content of a course or courses and curriculum is the philosophy, purposes, design, and implementation of a whole program. For White (1988), 'syllabus' can be easily referred to the content or subject matter of an individual subject. Dubin and Olshtain (1997) proposed that a syllabus is a more detailed and operational statement of teaching

---

<sup>1</sup> Email: mamra2006@yahoo.com



and learning elements which translates the philosophy of the curriculum into a series of planned steps leading towards more narrowly defined objectives at each level. But, in addition to these simple statements, there are definitions that pay more attention to the socio-constructive nature a syllabus may have. For Candlin (1984), syllabus is a social construction produced interdependently by teachers and learners concerning the specification and planning of what is to be learned. Hadley (1998) put steps further and held that a syllabus is an endorsement of a specific set of sociolinguistic and philosophical beliefs regarding power, education, and cognition ... that guide a teacher to structure his or her class in a particular way. Syllabus designing shapes the general framework of any educational course and is not only theoretically responsible for the content and materials of the course, but also for the outcome and practical considerations.

Regarding the syllabus, some classifications are suggested by scholars who focused on aspects of the term. White (1988) classified syllabi to *interventionist* (what should be learned) and *non-interventionist* (how the language is learned and integrated with learners' experiences). Long and Robinson (1998) Discussed about the '*Synthetic*' syllabi (which present L2 in a course of gradual accumulation of separately taught parts, and rely on the learner's ability to combine the pieces accurately) and '*Analytic*' syllabi (that present L2 in a process of natural sets of chunks, without linguistic control, and rely on the learner's competence to use the language appropriately). The distinguishing border between *product-oriented* and *process-oriented* syllabi is pioneered by Nunan (2000) in which the first concept focuses on 'knowledge' and the second; on 'learning experiences'. Following the dichotomy, Rabbini (2002) correspondingly emphasized the *ultimate specified goal* for product-orientation of language learning and its *flowing process* for the process-oriented look.

Syllabus design and materials preparation gives teacher a general and at the same time *specific* scheme for the education. As Nunan (2000, p. ix) stated, this general scheme "can be divided at least to three main critical branches: language knowledge with its conventional levels of linguistic description (skills & components), modes of behavior as the realization of the knowledge, and modes of action as a practical and applied rendering".

If we want to come at the body of any language learning syllabus such as functions and situations, tasks activities, language forms and communicative needs, various components can be pointed to. One specific and vitally important area with regard to the content of any syllabus is the realm of *Notions*; "the meanings a learner needs to express through language" (Richards & Schmidt, 2002, p. 365).

If language teaching aims to be genuinely professional and up-date, it seriously needs regular diagnosis, experimentation, evaluation and interpretation. In other words, the syllabus which is in



harmony with the *needs* of the learners, which does consider the persuasive forces and drives within the learners, and which really cares motivation, must be a pre-planned but not constantly-fixed one.

Imagine common books in language pedagogy as *Interchange*, *Headway*, *American English file* et al.. As can be seen, most of (in fact, all of!) what we have seen as syllabi for general English learning courses have been consisted of regular trends, routine patterns and specific frameworks. They introduce a series of forms, functions and situations and represent their corresponding materials within. Yes, this didn't happen at once. There exists a considerable body of research, practice and theorization a part of which are the outcomes like dividing syllabi based on their nature and function and describing them as tangible labels such as Structural/Grammatical, Functional/Notional, Synthetic/Analytic, etc. as the tangible labels.

There is a long history of interchanging ideologies and replacement of syllabus modes. But for the purpose of improving the educational system in terms of curriculum development and material Preparation, we need to do more. By having a glance at the regular general English books, we can see a variety of ideas, functions, communicative situations, and subjects raised and developed during the whole book. *Friends and Family*, *Food*, *Weather*, *Transportation*, *City*, *Marriage*, *Nature and Landscape*, *Health and Sanitation*, *Sports*, *Popular Culture* and so many other interesting issues are presented in the books with nuts-and-bolts of grammar, vocabulary and pronunciation (at the service of each other). This "theme-based/topic-based" system has really changed the condition from the age of merely syntactic and painfully structural patterns. In fact, the semi-revolution really helped many generations. But it seems that some other topics (somehow with a different nature) can be added to flourish the learning process. Those ideological topics which are suitable to be included in general English books can serve so many functions and if regulated well, can be so much fruitful.

This paper tries to introduce a new model for general English textbooks' design by inviting these ideological topics -which have always been criticized and accused of being irrelevant and even; destructive. Certainly, the idea of 'Critical Thinking' is of much relevance to the ultimate goal. Having discussed about the increasing attention to critical thinking in English language teaching, Pishghadam (2008) spoke about the recent emergence of various educational materials which focus on topics like 'justice', 'ideology' and 'politics' in language classes. He added that "proposition of critical theories by philosophers resulted to the tendency of teaching experts to critical thinking to the level that Piaget regards persons with high levels of critical thinking as the main goal of education in each country".



The paper argues that challenging topics will not deviate the learners, but in fact, stimulate their wisdom and flourish their judgment and outlook. In brief, they will open new horizons for the intellectual development of learners while they are engaged with language learning.

Usher and Edwards (1994) believed that Education must become the vehicle for the celebration of diversity, a space for different voices against the authoritative one. It may be true that embedding new (but evaluated) horizons to language pedagogy will lead to more effective teaching.

## 2. Theoretical Background

This article seeks if any change to our regular and cliché-type syllabus design principles and presuppositions can be changed in some way. Having reviewed the literature, we can claim that almost *nobody* has dealt with this critical issue, especially in its 'philosophical' form. In fact, throughout the bulky atmosphere of literature on topics as syllabus design, materials preparation, or curriculum development, at most, do exist reports in form of book, article, or project which have targeted the re-analysis of the syllabus design guidelines and parameters regarding the general 'process-oriented' -if not just 'circular'- considerations (the writers' googling on combination of *syllabus design + curriculum development + English teaching and learning + philosophy* had 288/000 resulted in no connection to the intended goal; they were just about the 'philosophy under language teaching/learning'). The reason behind can be the truth that their major problem has been the move on from the scholastic-typed programs and syllabus charts (*structural, grammatical, etc.*) to more dynamic, engaging, and fluid educational streams –or simply put; changing *product* to *process*.

According to our colleagues and TEFL fellows, the idea of embedding 'Philosophy' into language learning and teaching is not defined, yet, and its validity is under question. This unconsciousness, non-acceptance or even rejection is totally disputed by the underlying supposition of this article.

As mentioned, the two main reasons of such kind of resistance are educationally and morally, they can be that:

- 1- Most of us, English practitioners, think that philosophical topics are totally unrelated to educational goal-settings and principles. Hence; are keeping 'food-cloth-hobby' fashion in the main circle of syllabus design and wholly deny the educational (Human-making!) value of challenging topics as those in philosophy.



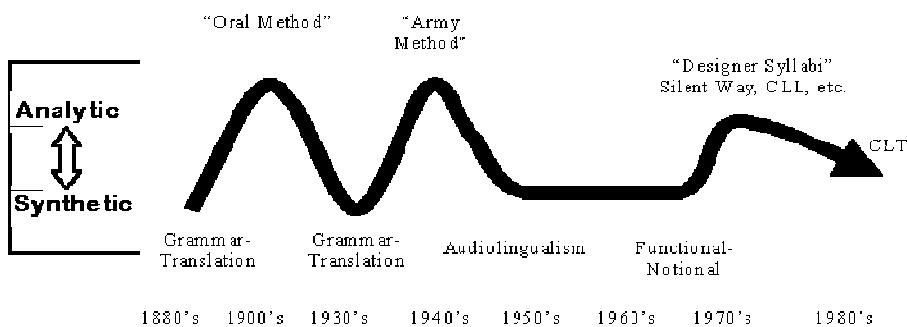
2- These challenging topics are kept intact due to the supposition that they are religiously or morally, critical; if not dangerous! Actually, it is mostly believed that including philosophical topics will not only be value-less, but can endanger the students' ideologies and beliefs.

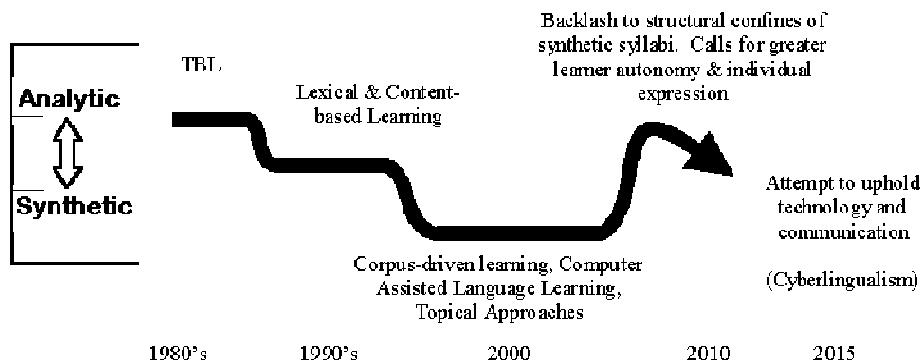
This article has persuasive responses to both of the raised issues and will invite teachers to include challenging themes and subjects during their educational setting.

### 3. The 'Fashionable' Taste!

Although all its contents can be seen as waves of a spectrum and not separated blocks, the history of syllabus design and curriculum development is in line with changes in the history of language teaching methodology (itself borrowed from Linguistic and Psychological movements). In late 18<sup>th</sup> and the beginning 19<sup>th</sup>, the faculty/scholastic phase, there was an intense force on learning writing and appreciating literature, composed of 'mastery' over grammar, translation, and essay-writing. The next phase (1930-1950) was the age of emphasizing structural patterns followed by embracing the 'practice' of aural/oral skills in order to 'overlearn'.

The cognitive era transformed the atmosphere to more 'mentalist' considerations engaging the cognitive and meta-cognitive trends to optimize the efficiency of learning. The next decades (1970s and 1980s) saw 'functional' individual-based moves which not only valued mind, but 'taste' and feeling. This was in harmony with the introduction of Communicative Competence (CC) by Dell Hymes (1972) which revolutionized the syllabi including that from then on, communicative needs were tried to be met, more (the story continues to the modernistic and post-modernistic reviews with counterparts as 'eclectic' and 'No-method talks!'. Hadley (2001) described this 'long and winding road' with a try to forecast the future trends.





**Figure1.**

Hadley's timeline of language teaching methodology (2001)

All mentioned ages have had their syllabus-related off-springs, of which the major paradigms contained a shift from product-orientation (with static and structural patterns) to process-orientation (with dynamic and non-linear frameworks). Our present situation in syllabus design is rooted in the so called *Communicative* era. When coming to recognize 'notions' and 'functions', we see that a lengthy list of general communicative themes and topics have been developed which helped learners in their language use in real settings. Titles as 'food', 'clothing', 'weather', 'sports', 'cinema', 'vacation', etc. and functions as 'introducing yourself', 'greetings', 'in a restaurant', 'asking price', 'asking addresses' and the like are actual manifestations of this communicative look which is undoubtedly necessary, ...but not enough! They are essential in guiding the learners in their personal expressions, social interactions and total engagement, but we must beware not to change them into a 'fashion' and think all is this! This article asserts that for two reasons, adding (and not; substituting) some challenging topics to the taste is required:

- 1- All human beings are full of important queries and serious questions regarding the radical epistemological issues. Educational settings has convenient atmosphere to answer the thirst based on intellectual atmosphere, mutual respect, and inherent reasoning beneath. In fact, the model doesn't abandon the 'Needs Analysis' of the learners, at all, but tries to put that if intellectual needs be embedded when students are supposed to learn language, they will be intellectually pushed forward and this leads to optimum and efficient learning.
- 2- Dealing with such kinds of issues is not tiring or disgusting for the students, and if regulated well, can highly stimulate their interest and motivation.



Like any framework, the tentative model going to be suggested by this paper needs careful regulation and adjustment.

#### 4. The Tentative Model

Firehammer (2007) talked about a reaction to the assumed certainty of scientific or objective efforts to explain the reality. After the modernistic revolution, in all the disciplines, this questioning stream is highly skeptical of the explanations which claimed to be valid for all groups, cultures, traditions, or races, and instead; focuses on the relative truths of each person. Education is not a separated area; nor is it far from the existing and changing atmosphere as in the postmodern age. It affects and is affected. For Beck (1993), postmodernism in education is a paradigm that challenges how we learn and appreciate knowledge in our lives. It questions the idea of a universal, unchanging, unified self or subject which has full knowledge of and control over what it thinks, says, and does.

The model presented in the article is certainly of a *postmodern* nature. According to Breen (1999), the pedagogy of the language classroom should deal with questioning absolutes, welcoming ambiguities, accepting uncertainties, participation in different and new discourses and exploring other identities. The main goal of the suggested model is to include philosophy and philosophical topics in English language teaching. The writers strongly believe that such embedding would not only do any harm to the syllabus, but also refresh the cliché-type framework by engaging the serious ideological concerns of the learners; hence, naturally motivate them to both seek their answers and learn the language lesson.

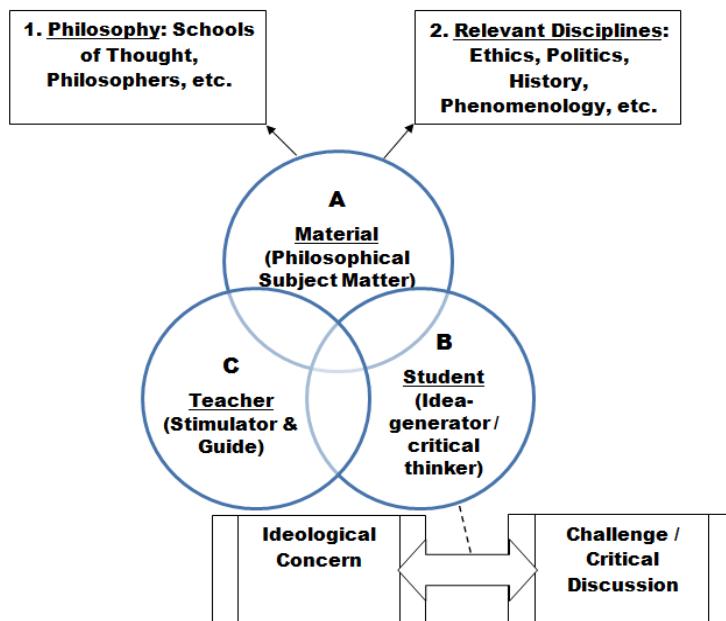
By the umbrella term 'Philosophical Issues', we mean all those themes, topics, subject areas and fields which deal with our serious ideological concerns. By implementing such a view, certainly, terms like 'Cosmology', 'History', 'Epistemology', 'Existentialism' and 'Ontology' are in the ground. In fact, in addition to the educational targets indicated, the underlying goal is to if not answer some serious intellectual questions, at least creating the needed concern and stimulating the potential wisdom in relation to addressing them in the main thematic syllabus of language teaching.

Surely, before coming at such a plan, there are many things to be considered. Here are some:

- 1- First is the suitability of this inclusion and whether it is a deviation from the educational goals defined.
  - The writers believe that including philosophical topics cannot -by any means- considered to be a 'deviation' from the syllabus because these topics, as those *communicatives* presented now are of concern to students; if not more important for them.



- 2- Second is the appropriateness of challenging issues for learners at all levels.
- Well, it is quite clear that the mental ability and/or interest of learners can differ greatly. They may not be intellectually; *homogeneous*. The main concern of the writers was to introduce/maneuver on such topics in relation to upper-intermediate and advanced learners in terms of English level. But in terms of age, although not intended; chiefly, there exist simplified materials for the beginners or even children. Titles such as 'Philosophy in Simple Language', 'Philosophy for Beginners' or 'Philosophy for Children' have been among the best sellers. They introduce philosophical topics in a user-friendly, lovely and cute manner which attracts the readers. Selection, gradation and modification of appropriate materials from these sources can help the teacher much, if s/he intends to apply the model for low-level learners.
- 3- Third: appropriateness of including all philosophical topics in language teaching syllabus.
- There is no need (neither is it valuable or advantageous) to include all topics and subjects discussed in philosophy, even in their simplified versions. To tell the truth, it actually damages the whole building! What we try to introduce is a balanced 'sipping' of the area which both creates interest in students and answers some of their concerns, as well.



**Figure2.**

Representation of the tentative model (Philosophical Syllabus) with its building blocks totally constitute a curriculum of the program



#### *4.1. Structural organization*

As far as the 'label' is concerned, the model proposed by this article goes under the 'content' syllabus but the content is not 'communicative' in its fashionable meaning. In fact, to invent a wheel, it can add 'challenging' syllabus, to the literature, theoretically. The usual process of selection and gradation of the materials is done, in this syllabus, too. The content is best tuned if it is arranged according to the simplicity and harmonized based on shared features (whether some topics go under one title, hypothetically). Richards and Renandya (2002) classified the curriculum development models to 'Classical Humanism' (structural/grammatical), 'Reconstructionism' (notional/functional) and 'Progressivism' (process/procedural). The proposed model in this paper goes somewhere between the Reconstructionism and Progressivism and values both 'objective' and 'process'.

#### *4.2. Content*

Probably, the central issue concerning the syllabus design is the problem of choosing the appropriate subject matter. This article recommends if Philosophy is a good one! The Encyclopedia Britannica (2012) defines 'Philosophy' as "the critical examination of the grounds for fundamental beliefs and an analysis of the basic concepts employed in the expression of such beliefs".

Since the aim is to introduce learners with some fundamental approaches to beliefs; critically, the main building blocks of the syllabus content is made up of issues directly related to philosophy. These include: History, Schools of Thought, Religion, Metaphysics, Philosophy of Language, Philosophy of Mind, Existentialism, Idealism, Skepticism, Phenomenology, Law, Politics, Ethics, Arts and Aesthetics, and so forth.

It's noteworthy to mention that this doesn't mean that a whole unit of a lesson be dedicated to e.g. Western Philosophy. These are the *themes* and *motifs*. The actual unit of teaching can be materials and discussions beginning about preliminary philosophical questions. The main body can be attributed to analysis of the history of human thought: major schools of thought with their representative thinkers and philosophers. The content can be arranged from Greek philosophy, moving across the Hellenistic and Medieval, and entering the Modern age.

As mentioned, the content should be simplified (lexically and grammatically) and made user-friendly based on the level of the intended students. The learners must at the same time comprehend the texts and see questions raised from within based on those comprehensions which are based on getting familiar with stimulating concepts, notions and problematic areas. To give a clear picture of how this syllabus looks like, table below is presented in terms of different parts and their sequence



divided based on units (the table is set for an imaginary 20-session course. It can, surely, be limited or expanded according to the necessities of the class, with careful regulation).

**Table 1.**

General framework for the syllabus

<u>Part A</u>	<u>Preliminary Concepts</u>		<u>Main Question</u>
<b>Unit 1</b>	Existence		- What is the nature and qualities of the Existence?
<b>Unit 2</b>	God		- Is there a creator? - What is his essence & characteristics?
<b>Unit 3</b>	Nature		- How is the relation of human with nature? How does nature affect our life?
<b>Unit 4</b>	Determinism/Free Will		- Are we doomed to be puppets or we have choice, ourselves? How is the extent of our free will?
<b>Unit 5</b>	Religion		- What is Religion? - Is it necessary to follow a religion?
<b>Part B</b>	<u>Philosophical Thought</u>		
	<u>Major Figures</u>	<u>Schools</u>	<u>Main Question</u>
<b>Unit 6</b>	Socrates	Socratic Age	- Can we avoid the dangers of relativism, which seems to undermine morality and make the pursuit of truth impossible?
<b>Unit 7</b>	Plato	Platonic Age	- Mustn't we be committed to some more eternal and unchanging ideals if we are going to be committed to goodness and truth?
<b>Unit 8</b>	Aristotle	Peripatetics	- Can't we get a clear grip on knowledge and goodness, without being committed to some unrealistic ideals which we can't experience?



<b>Unit 9</b>	John Locke & David Hume	Empiricism	- Given that sense experience is our only source of knowledge, how far can knowledge extend, and what are the inevitable limitations?
<b>Unit 10</b>	René Descartes & Immanuel Kant	Rationalism & Epistemology	- Given that reason is our only reliable source of knowledge, what can we deduce about reality from pure thought, and how far can we trust the appearances of sense experience?
<b>Unit 11</b>	Søren Kierkegaard, Friedrich Nietzsche, & Jean- Paul Sartre	Existentialism	- If we accept our feeling of mental freedom as being true, how should we exercise this responsibility in our lives?
<b>Unit 12</b>	Jacques Derrida	Post-Modernism	- What follows from the fact that relativism is right, and truth and morality change continually with culture and prejudice?
<b>Unit 13</b>	Confucius, Laozi & Buddha	Eastern Philosophy: Chinese & Indian/ (Daoism, Buddhism)	- How is the quality of God, Creation of the world, & Life? - How can we escape from the sufferings of life and from the endless reincarnation caused by human desires?
<b>Unit 14</b>	Al-Farabi, Avicenna, Mulla- Sadra,	Islamic Philosophy	- Can we search for Hekma (wisdom) and save ourselves? - How can Kalam (Islamic theology through Dialectic) help us?
<b>Part C</b>	<b>Relevant Trends</b>		<b>Main Question / Problem</b>
	<b>Discipline</b>	<b>Major Figures</b>	



<b>Unit 15</b>	Cosmology, Metaphysics & Ontology	Saint Thomas Aquinas	- The nature of existence & ultimate reality
<b>Unit 16</b>	Ethics	Jeremy Bentham	- Correct actions are those that result in the greatest happiness for the greatest number of people.
<b>Unit 17</b>	Politics	Jürgen Habermas, Karl Marx	- Is modern scientific knowledge and research; objective and value-free? - The working class should rebel and build a communist society.
<b>Unit 18</b>	History	Georg Wilhelm Friedrich Hegel	- Truth is reached by a continuing dialectic, in which a concept (thesis) always gives rise to its opposite (antithesis), and the interaction between these two leads to the creation of a new concept (synthesis).
<b>Unit 19</b>	Philosophy of Art & Aesthetics	Walter Benjamin & Jean François Lyotard	- Essence and perception of beauty and ugliness & whether such qualities are objectively present in the things they appear to qualify, or whether they exist only in the mind of the individual
<b>Unit 20</b>	Phenomenology	Martin Heidegger	- The structures of experience as they present themselves to consciousness, without recourse to theory, deduction, or assumptions from other disciplines such as the natural sciences

As can be seen, the whole content syllabus is divided to three parts starting from stimulating sparkles about nature and existence (*Preliminary Concepts*), moving through the following reasonings (namely; *Philosophical Thought*), and finally proposing some special disciplines resulted from philosophy (*Relevant Trends*). This content can be taught in one, two or even three semesters, according to the level of the students.



At first, any TEFL practitioner may wonder if we are intending English for Specific Purposes (ESP) curriculum, actually. But this is not true, because in ESP course, the major goal is teaching that intended non-linguistic subject matter (physics, mathematics, etc.) through the medium of language and familiarity with the jargon and technical terms (words, expressions, and even structural patterns) but this is not so, here. When we propose this *philosophical* syllabus, the main aim is to teach language to students but the subject matter is at service of philosophy (because of the belief that this branch of science is a radical one and acts completely different from technical and academic disciplines. Hence, is highly stimulating).

Perhaps the Needs Analysis (NA) background of the model is best described by referring to Brindley (1989, p. 64) who postulated that there are two orientations regarding the needs-analysis: a narrow 'product-oriented' view and a broad 'process-oriented' one. The process-oriented view of needs take into account factors such as learner motivation; the one which our model seeks.

#### *4.3. Parallel activities*

As far as the challenging issues and implementing a critical look are concerned, having a structured material even with numerous eye-catching topics is not enough! There is crucial need for discussion challenge in the classroom so that the theoretical concerns are practically put into the practice of debate. In such a curriculum, what Prabhu (1987) entitled '*procedural*' syllabus, Breen (1987) named '*process*' syllabus, and Crookes (1992) labeled '*task-based language teaching*' are not irrelevant if added to philosophical/ideological seasonings. Long and Crookes (1992) hold that central to the procedure syllabus is the belief that knowledge of linguistic structure develops largely subconsciously through the opinion of some internal system of abstract rules and principles on the bases of extensive input from the target language. This comes about when the focus is on the meaning, and the condition is best met when trying to complete tasks.

Moreover, the teacher has an important role. Pishghadam (2008) proposed some suggestions to teachers for growth and development of critical thinking in language class and having really 'critical' thinkers.

- 1- Letting students express their attitudes, freely
- 2- Not trying to 'impose' any idea in class
- 3- Making learners question themselves
- 4- Guiding them in how to reach a valid question and pursue its answer
- 5- Creating a low-anxiety and enjoying atmosphere



In addition to the above-mentioned guidelines, teachers can make students question different subjects and phenomena and even themselves. They should abandon the cliché of 'description' for 'explanation'. Arriving at definite clear-cut answers is not the goal, at all, but questioning existence, quality and several aspects of actions and reactions have great importance.

## 5. Discussion

The inclusion of non-linguistic objectives in language learning syllabus is not a new or unknown issue. Richards (2001) emphasized on 'non-language' outcomes of language learning class, which include affect cultivation (such as confidence, motivation, and interest), learning strategies, thinking skills, interpersonal skills, and cultural understanding. The underlying assumption of this trend in syllabus design is that, as a school subject, language education should not merely aim at helping students learn language knowledge and skills. Rather, it has responsibility in fostering students' *whole-person* development.

Usher and Edwards (1996) proposed that education is ambiguous. It both seeks and rejects closure. It is both *closed* and *open*. It can be an instrument/a device for control and legitimization, but also it has the potentiality to question the status of the definitive, the certain and the 'proven'. Pishghadam (2008) asserted that for experts, critical thinking is one of the most important educational principles in each country, and that every society needs critical thinkers to grow and flourish. This fact can best be achieved through having a -this time- 'major' look at the philosophy in our general English classes probably because the general root of critical theories can be found in philosophers' thought (Rasmussen, 1996).

In addition to 'inviting philosophy to the scene', the nature of what the presented model recommends has considerable looks on terms as *Process-orientation*, *Nonlinearity*, *Educational Ecosystem*, *Dynamics*, and *Fuzziness*. It is believed that triggering the basic beliefs and concerns and attending to fundamental ideologies will naturally increase the intrinsic motivation of the students. This curiosity-raising is, here, experimented in the English teaching context. The model embraces philosophy and holds that related subject matters can open new horizons for the intellectual development of learners while they are engaged with language learning. It must be said that the main educational/ideational intentions are to shape a kind of creative thinking style, inviting students to *think* about the issues -in addition to their usual language-related practices-, and helping them move toward a more self-centered and responsive individuality.

Slattery (2006) emphasized that the postmodern curriculum should promote a creative search for deeper understanding through interdisciplinary and inclusive tasks, projects and narratives and that



its content should problematize, interrogate, contextualize, and challenge the learners. The *philosophical syllabus* suggested in this article tries to be in harmony those objectives.

## References

- Beck, C. (1993). *Postmodernism, pedagogy, and philosophy of education*. Retrieved on June 9, 2012, from [http://www.ed.uiuc.edu/eps/PES-Yearbook/93\\_docs/BECK.HTM](http://www.ed.uiuc.edu/eps/PES-Yearbook/93_docs/BECK.HTM)
- Breen, M. (1999). Teaching language in the postmodern classroom. *Barcelona English Language and Literature Studies (BELLS)*, 10, 47-64.
- Brindley, G. P. (1989). General-purpose language teaching: A variable focus approach. In C. J. Brumfit (Ed.). *General English syllabus design*. ELT Documents No. 118. London: Pergamon Press & The British Council. 61-74.
- Candlin, C. N. (1984). Syllabus design as a critical process. In C.J. Brumfit (Ed.). *General English syllabus design*. ELT Documents No. 118. London: Pergamon Press & The British Council. 29-46.
- Dubin, F., & Olshtain, E. (1997). *Course design: Developing programs and materials for language learning*. Cambridge: Cambridge University Press.
- Firehammer, R. (2007). *Postmodernism*. Retrieved from <http://www.spaceandmotion.com/Philosophy-Postmodernism.htm>
- Graves, K. (Ed.). (1996). Teachers as course developers. New York: Cambridge University Press.
- Hadley, G. (2001). Looking back and looking ahead: A forecast for the early 21st century. *The Language Teacher*, 25(7), 18-22. Retrieved on June 12, 2012, from <http://nus.ac.jp/~hadley/publication/21centforst/forecast.htm>
- Long, M., & Crookes, G. (1992). Three approaches to task-based syllabus design. *TESOL Quarterly*, 26, 27-56.
- Long, M. H., & Robinson, P. (1998). *Focus on form: Theory, research and practice*. In Doughty, C. J., and Williams, J. (Eds.), *Focus on form in second language acquisition* (pp. 15-41). Cambridge: Cambridge University Press.
- Nunan, D. (2000). *Syllabus design*. UK: Oxford University Press (OUP).
- 'Philosophy'. (2012). Encyclopædia Britannica. *Ultimate reference suite*. Chicago: Encyclopædia Britannica.
- Pishghadam, R. (2008). Afzayesh-e tafakor-e enteghami az tarigh-e mobahese adabi dar kelashaye zaban-e Englisi. [Enhancing Critical Thinking Through Literary Discussion in English Language Classes]. *Majale-ye Daneshkade-ye Adabiat va Olum-e Ensani-e Mashhad*, 159, 153-170.
- Rabbini, R. (2002). An introduction to syllabus design and evaluation. *The Internet TESL Journal*, 8(5). Retrieved on May 8, 2012, from <http://iteslj.org/Articles/Rabbini-Syllabus.html>



- Rasmussen, D. M. (1996). Critical theory and philosophy. In Rasmussen, D. A. (Ed.), *The handbook of critical theory* (pp. 11-38). Oxford: Blackwell.
- Richards, J. C. (2001). *Curriculum development in language teaching*. Cambridge: Cambridge University Press.
- Richards, J. C., & Schmidt, R. (2002). *Longman dictionary of language teaching and applied linguistics*. UK: Longman.
- Richards, J. C., & Renandya, W. A. (2002). *Methodology in language teaching: An anthology of current practice*. UK: Cambridge University Press.
- Slattery, P. (2006). *Curriculum development in the postmodern era*. New York: Routledge Taylor and Francis.
- Usher, R., & Edwards, R. (1994). *Postmodernism and education*. Routledge, London.
- White, R. (1988). *The ELT curriculum, design, innovation and management*. Oxford: Basil Blackwell.

### Bio Data

**Abbas Moradan** holds a Ph.D. in TEFL and is a faculty member at the Department of English Language and Literature at Semnan University. His research interests include Material Preparation, Language Testing, Translation and Literature.

**Mohammad Amerian** holds a B.A. in English Language and Literature and an M.A. in Teaching English as a Foreign Language (TEFL) from Semnan University, Iran. He has taught English in various levels. Amerian's research interests center around Psycholinguistics, Social Psychology, Semiotics, Language and Socio-Cultural Studies and Pragmatics.



## Improving Complexity, Accuracy and Fluency in EFL Learners' Oral Production through Computer-Mediated Emotional Intelligence Activities

Parisa Abdolrezapour<sup>1</sup>

University of Isfahan

### Abstract

This study aims to gain more in-depth understanding of the possibility of enhancing EI through computer-mediated emotional activities and to see whether it is possible to improve EFL learners' oral production measured in terms of complexity, accuracy and fluency through exposing them to computer-mediated emotional intelligence activities. The participants of the study consisted of 33 students studying Interchange 2 course in a language center in Iran. Participants in the experimental group received EI training, in the form of some movies with highly emotional content with the express purpose of inducing them to talk about their emotions and raising their emotional intelligence. However, in the control group, participants were taught under institute's normal procedures that did not place any emphasis on activities with emotional content. Results showed a significant increase in EI scores of the experimental group and no significant change was observed in the control group's EI scores. Moreover, learners of the experimental group were able to produce more accurate and fluent language than the control group. The difference between these two groups in terms of complexity measures, however, did not reach significance.

**Keywords:** Accuracy, Complexity, Emotional intelligence, Technology, Oral fluency

### 1. Introduction

For decades, schools have tried to predict which students would do well both in higher education and in the workplace. The tools they used to make their best guess were standardized achievement tests and IQ scores. Recent studies, however, indicate that there may be a better predictor--the measurement of a child's emotional intelligence (Bar-on & Parker, 2000). Experts now believe that success is influenced 20 percent by IQ and 80 percent by various factors that constitute a person's character and personality, or their *emotional intelligence* (henceforth, EI). Therefore, if teachers are

---

<sup>1</sup> Email: abdolrezapour@gmail.com



concerned with helping students develop their L2 abilities, they might want not only to improve their EI skills but also reduce foreign language anxiety, if possible.

The term ‘Emotional intelligence’ was first used in the pioneering work of Salovey and Mayer (1990). They have defined emotional intelligence by the specific competencies it encompasses, “the ability to perceive emotions, to access and generate emotions so as to assist thought, to understand emotions and emotional knowledge, and to reflectively regulate emotions so as to promote emotional and intellectual growth” (Mayer, Salovey, & Caruso, 2004, p. 197). The two faculties related to socio-emotional competence – i.e. *interpersonal* and *intrapersonal* intelligences – as proposed by Gardner (1993) – resemble Salovey and Mayer’s distinction made between processes regarding one’s own emotions and the emotions of others. Later, Daniel Goleman brought it into the academic performance mainstream with the publication of his 1995 book entitled “*Emotional intelligence: Why it can matter more than IQ*”. According to him, EI is a combination of five domains: a) knowing one’s emotion, b) managing one’s emotion (i.e. handling fear, anxiety), c) motivating oneself (emotional control, the ability to delay gratification), d) recognizing emotions in others, and e) handling relationships. Following Goleman (1995), a number of studies showed that social and emotional skills play a central role in one’s academic, social and personal life above and beyond his general intelligence (Parker et al., 2004; Song et al., 2010).

According to Lynn (2002), emotional intelligence explains why, despite equal intellectual capacity, training or experience, some people excel while others of the same caliber lag behind. Among those studies which pay attention to the effect of emotional intelligence on academic success in education is the one carried out by Stottlemayer (2002). In a study of EQ and its relation to student achievement among 200 eleventh and twelfth grade American students in Texas, Stottlemayer found that EI skills were statically significant predictors of academic achievement. Rossiter (2003) reports differential success in second or foreign language learning as being attributed to individual differences such as intelligence, aptitude, personality, motivation, attitude and anxiety. More recently, Downey, Mountstephen, Lloyd, Hansen and Stough (2008) found that high EI contributes to increased motivation, planning, and decision making, which positively influence academic performance.

Having been concerned about the relation between EI and one’s success, the possibility of enhancing one’s EI has been discussed in a number of studies (e.g. Hunt & Evans, 2004; Richardson, 2002; Schutte et al., 2001). Schutte and her colleagues (2001) discussed that though researchers have conceptualized EI as relatively stable, it might be possible to increase it through some kinds of intervention. In the same line, Richardson (2002) claimed that emotional intelligence can be nurtured in young adolescents by teaching them components of EI such as coping skills,



how to acquire and use information, how to work with others and how to manage personal growth. According to Hunt and Evans (2004), EQ can be increased with tuition; thus, it is not a fixed ability.

In L2 teaching literature, one of the first studies addressing the relationship between emotional intelligence and second/foreign language learning was Aki (2006). He considered EI, i.e. having the ability to recognize, employ, comprehend and manage emotions, more important in language learning than possessing high intelligence values and his findings lend support to the positive effects of enhancing emotional intelligence on language acquisition, where the enhancement of emotional intelligence was operationalized as stimulating the imagination, humor, and creativity in young language learners. In another attempt, Dewaele Petrides and Furnham (2008) found a negative relationship between trait EI and foreign language anxiety (FLA) which was defined as "the feeling of tension and apprehension specifically associated with second language (L2) contexts, including speaking, listening, and learning" (MacIntyre & Gardner, 1994, p. 284). More recently, Abdolreza and Tavakoli (2012) found a high positive correlation between one's achievement in reading comprehension and her/his emotional intelligence. In addition, greater achievement in reading comprehension was found for those who had been through literature response activities which were used to foster their emotional intelligence.

The "study of applications of the computer in language teaching and learning" is called computer assisted language learning (CALL) (Levy, 1997, p. 1) and it can be traced back to as early as 1960s. In this approach computer can be used as a facilitative tool in all phases of language learning and teaching – from the very first stages of presentation and practice to the final stages of assessment and evaluation. Studies on the language-acquisition value of computer-mediated tasks have proliferated since the early 1990s (Beauvois, 1992; Kern, 1995; Smith, 2005). A glance at the topic archive of the journal of language learning and technology shows that previous studies have addressed distance learning (Liang, 2010), collaborative learning and social context provided (Elolah & Oskoz, 2010) with a focus on different language skills (reading, writing, listening, speaking, vocabulary and grammar). However, the role of emotions and motivation, and more specifically, the role of emotional intelligence in computer-mediated foreign language learning has been a slightly neglected area of enquiry.

Nevertheless, some attempts tried to find the possibility of applying technology to enhance EI in other contexts. The possibility of developing interpersonal skills was investigated in Holsbrink-Engels's (1997) study through what he called *computer-based role-playing*. And Goldsworthy's (2003) study provides a comprehensive overview of the literature of EI-development via technology. He argues that technology can be used to foster learners' social skills and social problem solving as they are engaged in tasks at the computer and he further provides a framework for design and



development of technology-based instruction for emotional intelligence. His STAR (standing for *stop, think, act, reflect*) project is a project-solving model addressing several key areas of EI framework. Therefore, there seems to be good reason to believe that computer-mediated emotional activities may be appropriate tools for raising learners' emotional intelligence.

This study aims to gain more in-depth understanding of the possibility of enhancing EI through computer-mediated emotional activities and to see whether it is possible to flourish EFL learners' oral production measured in terms of complexity, accuracy and fluency through exposing them to computer-mediated emotional intelligence activities.

## 2. Method

### 2.1. Participants

The population under study included the EFL learners who enrolled in a language center in Iran. A total of thirty three students studying Interchange 2 participated in this study. Learners were native speakers of Farsi and they had taken English courses for three to four years. None of these participants had additional exposure to the English language, apart from the normal TV programs and the Internet, which are almost impossible to control for. Our participants were in two intact classes taught under same instructor. Their level of English proficiency is intermediate; that is, all participants are currently taking the intermediate EFL course offered by the institution and their language proficiency was tested by the grammar part of the 'Oxford Placement Test 2' (Allan, 1992). One class was taken as the experimental group with 18 students and the other as the control group with 15 students. All participants were female and they varied in age from 15 to 18.

### 2.2. Instruments

#### 2.2.1. Measures of oral performance

Most SLA researchers agree that facets of L2 oral performance are themselves multifaceted and entail several sub-constructs; therefore there is a need for using multiple measures for assessing each construct (Housen & Kuiken, 2009). Nevertheless to avoid what Norris and Ortega (2009) call redundancy in measurement, researchers are recommended to use complementary but distinct measures for assessing each principal construct. In the present research, drawing on the host of studies into the CAF triad (Ellis & Yuan, 2004; Yuan & Ellis, 2003) the following measures were used to assess three major dimensions of second language oral performance:



Complexity measures:

- Syntactic complexity (amount of subordination): the ratio of clauses to AS-units (the Analysis of Speech Unit) in the participants' production. Following Foster, Tonkyn and Wigglesworth (2000), the AS unit, defined as "... a single speaker's utterance consisting of an independent clause or sub-clausal unit, together with any subordinate clause(s) associated with it" (Foster, et al., 2000), was used as a unit for measuring syntactic complexity.
- Syntactic variety: the total number of different grammatical verb forms used in participants performances. I used tense and modality as grammatical verb forms for the analysis.

Accuracy measures:

- Error-free clauses: the percentage of the clauses which were not erroneous. Error-free clauses were defined as clauses in which no error was seen with regard to syntax, morphology, native-like lexical choice or word order.
- Correct verb forms: the percentage of all verbs which were used correctly in terms of tense, aspect, modality, and subject-verb agreement.

Fluency measures:

- Rate A (number of syllables produced per minute of speech): the number of syllables within each narrative, divided by the number of seconds used to complete the task and multiplied by 60.
- Rate B (number of meaningful syllables per minute of speech): Rate A's procedure was followed again, but all syllables, words, phrases that were repeated, reformulated, or replaced excluded.

#### *2.2.2. Trait Emotional Intelligence Questionnaire-Adolescent Short Form (TEIQue-ASF)*

In this study, the short form of the TEIQue (Petrides, Sangareau, Furnham, & Frederickson, 2006) comprised of 30 items was used to measure the students' emotional intelligence prior to the experiment to make sure of their initial comparability. TEIQue-ASF is a simplified version, in terms of wording and syntactic complexity, of the adult short form of the TEIQue developed for use with adolescents aged 12-18 years. All items are sampled from the 15 subscales of the adult trait EI sampling domain (two items per subscale). The test yields scores on four factors, namely well-being



(covering self-esteem, happiness, and optimism), self-control (covering low impulsiveness, stress management, and emotion regulation), emotionality (covering emotion expression, relationships, empathy, and emotion perception), and sociability (covering assertiveness, emotion management, and social awareness) in addition to global trait EI. Example items include 'I can control my anger when I want to', 'I'm happy with my life', and 'I'm good at getting along with my classmates'. Higher scores on the TEIQue-ASF indicated higher levels of trait EI. We opted for the short version with 10-minute completion time, because we had time limitation and there was a concern that the participants might not be able to complete the longer version (e.g. due to reading difficulties). Subjects responded on a 7-point Likert scale continuum from 'Completely Disagree (number 1)' to 'Completely Agree (number 7)'.

### *2.3. Procedures*

To ensure subjects' homogeneity prior to conducting the study, they were asked to answer TEIQue-ASF (Petrides et al. 2006) and an oral narrative task was employed to elicit oral language performance. In this task a sequenced set of picture prompts were shown to the participants while they were asked to narrate the story. Then, learners went through the eight-week treatment. Participants in the experimental group received EI training, based on Goleman's EI framework, in the form of some movies with highly emotional content followed by some texts and images in which they were asked to work individually (when they had to provide the answer for the computer by typing or clicking on an item) or collaboratively (when they had to find the answer by discussing it in a group or when they were asked to work on interpersonal tasks such as finding emotions in others) to find answers which mainly involved emotional background.

#### Goleman's EI framework:

- 1. Knowing one's emotions:** to make learners aware of their emotions, the program re-showed them some parts of the movie in which the character was feeling an emotion and asked the learner to find what caused that particular emotion.
- 2. Recognizing emotions in others:** to improve learners' level of empathy, the program encouraged them to pay attention to the body language of characters of the movie. In the final stage, parts of the movie which could help them gain a better understanding of others' emotions were displayed again.
- 3. Managing one's emotions:** to help learners learn how to manage their emotions, there was a pause on a part of movie where the character felt an emotion (e.g. sadness or anger) and then learners were asked to find the cause of that emotion (i.e. what triggered it).
- 4. Handling relationships:** to make students aware of the forth characteristics of Goleman's theory, learners were invited to watch the parts of the movie which showed interactions



between characters again. Then, they were asked to think about themselves in characters' positions.

**5. Motivating oneself:** to raise learners' level of self-motivation, some parts of the movie that characters experienced a positive feeling were shown again and they were asked to talk about the ways characters could change their negative feelings to positive ones.

It was assumed that performing such activities would enable learners to become increasingly aware of their emotional states. As said above, movies shown to the subjects in the control group were void of emotional words and content and learners were taught under the ordinary procedures in which no explicit emphasis was placed on emotions during oral communication activities.

The data obtained were ultimately analyzed using the statistical package for social sciences (SPSS) to provide plausible answers to the research questions posed above. First, descriptive statistics were used and the obtained scores were checked in terms of the normality of distribution using such indices as Kurtosis and Skewness. Then, independent sample t-tests and Pearson Product correlation procedures were performed.

### 3. Results

Table 1 shows the descriptive statistics and the results of t-tests for the scores obtained from subjects' TEIQue-ASF and their performances on the oral narrative task on pre-tests. As is shown in the table, the experimental and control groups were fairly similar in terms of all measures prior to the treatment and results of the independent-sample t-tests confirm that there were no statistically significant differences between the two groups in terms of TEIQue-ASF or task performance measures ( $p > 0.05$ ).



**Table 1.**

Descriptive and inferential statistics on pre-tests

Variables	Experimental Group (N=18)		Control Group (N=15)		Sig
	Mean	SD	Mean	SD	
TEIQUE-ASF	116.78	7.67	119.05	9.75	.179
Syntactic Complexity	.651	.11	.627	.07	.128
Syntactic Variety	3.98	.27	4.21	.32	
Percentage of correct verb form	43.12	4.2	44.19	5.1	.237
Percentage of error free clauses	35.68	2.11	37.88	3.01	
Number of syllables produced per minute of speech	33.42	3.2	31.21	2.8	.325
Number of meaningful syllables per minute of speech	21.25	2.7	22.05	2.4	

Level of significance is 0.05

In order to see whether the positive effects of exposing learners to emotional activities transferred to their EI and oral performance measures, the mean scores were compared across the two groups after the two months period. Results of the inferential statistics displayed in Table 2 point to the statistically significant differences between the experimental and control groups in terms of EI and task performance measures ( $p < 0.05$ ).



**Table 2.**

Descriptive and inferential statistics on post-tests

Variables	Experimental Group (N=18)		Control Group (N=15)		Sig
	Mean	SD	Mean	SD	
TEIQUE-ASF	142.38	9.93	121.75	8.25	.000*
Syntactic Complexity	.712	.10	.698	.11	.09
Syntactic Variety	4.25	.18	4.31	.27	
Percentage of correct verb form	49.12	3.94	46.29	4.6	.03*
Percentage of error free clauses	40.78	1.09	38.18	2.71	
Number of syllables produced per minute of speech	37.52	2.8	33.11	1.9	.00*
Number of meaningful syllables per minute of speech	28.37	2.4	29.11	2.7	

\*Level of significance is 0.05

Results showed a significant increase in EI scores of the experimental group and no significant change was observed in the control group's EI scores. Moreover, learners of the experimental group were able to produce more accurate and fluent language than the control group. The difference between these two groups in terms of complexity measures, however, did not reach significance.

#### 4. Discussion

The aims of this study were: (a) to investigate the possibility of enhancing learners' emotional intelligence through computer-mediated emotional activities; and (b) to find out whether there is a significant difference between the oral performance (measured in terms of complexity, accuracy and fluency) of those exposed to emotional activities and those instructed under the institute's normal procedures. The observed significant differences between the EI scores of the experimental and control groups could be attributed to the influence of the technology-mediated activities with its



numerous opportunities for direct exposure to emotional activities. This finding, which is consistent with our prediction, is also in accord with the arguments put forth in the literature vis-à-vis the possibility of raising one's EI (Hunt & Evans, 2004; Richardson, 2002; Schutte et al., 2001) and those who discussed the possibility of raising EI via technology (e.g. Goldsworthy, 2003; Holsbrink-Engels, 1997)

A number of studies have documented empirical evidence in support of the positive relationships between EI and academic success (Parker et al., 2004; Song et al., 2010) and there is some evidence indicating that EI and second language performance are positively related (Abdolrezapour & Tavakoli, 2012; Aki, 2006; Dewaele et al., 2008). In this way, it can be said that the observed increase in learners' EI is beneficial for their language performance and as the results from the task performance measures showed that there is a significant difference between the control group and those who underwent emotional activities in terms of accuracy and fluency. A plausible interpretation for this difference is that participants in the experimental group, through their extensive exposure to emotional activities, became more motivated (based on Goleman's (1995) model motivation is one of the five characteristics that constitute emotional intelligence) and this motivation, in turn, positively influenced their academic performance. This finding is consistent with our prediction and provides further empirical support for Abdolrezapour and Tavakoli's (2012) findings.

The intervention proposed here involves far more than promoting one's speaking ability; in effect, it has the potential to help learners become more socially competent and offers them access to a range of knowledge and abilities which might enable them to have better performance in academic context and in their social life. This, however, cannot be supported with the data obtained for the present study and thus needs to be tested in further research.

Interestingly enough, the experiment was appealing to the students. When the instructor explained to these students what the purpose of the study was -namely, to discover what might be the effect of dealing with emotions on their language performance- they were all enthusiastic about and interested in taking part in it. In addition to understanding that their participation would provide valuable information for teachers of second language, they were excited by such notions as emotional intelligence and affective factors. They had never thought about their emotions, at this level of awareness, and the possibility of its effect on their language learning. As one student put it, "I can't believe we're going to talk about our emotions in the class. Nobody ever did that before in our classes".



All in all, this study aimed at improving our understanding of the relationship between emotional intelligence and academic performance in general as recommended by Parker et al. (2004) and Song et al. (2010) and the acquisition of one specific aspect of language –as recommended by Abdolrezapour and Tavakoli (2012). Moreover, the role of technologies as a tremendously valuable tool to enhance one's EI was scrutinized. The findings point to substantial differences among the two participatory groups (experimental and control groups) both in EI scores and in speaking proficiency, suggesting that it is possible to enhance one's EI in an EFL setting. Moreover, it is presumed that existing differences in their oral performance reflect the efficiency of emotional program in improving learners' oral ability and the possible correlation between emotional intelligence and task performance in an EFL context.

As for the pedagogical implications, this study provides evidence that emotions can be successfully implemented in EFL classrooms for young beginners and are at least as effective as learners' speaking ability is concerned. Hence, one implication of this study is that tasks designed for EFL classrooms might address the needs and preferences of certain students, taking into account learners' individual differences and more importantly their emotions resulting in more motivating tasks.

This study, however, is not without limitations which could impede the generalization of the results to other situations. First and foremost, if a larger sample size were considered, the obtained results could have been more significant and reasonable than they are now. Second, although the implementation of EI approach in an EFL classroom may prove to be incentive and useful for students, teachers may lag behind so as to terminate the formal syllabus in due time. Finally, the practical problem is that, to appropriately use such a procedure in the classroom requires more competent and experienced instructors that are almost always unavailable.

## References

- Abdolrezapour, P., & Tavakoli, M. (2012). The relationship between emotional intelligence and EFL learners' achievement in reading comprehension. *Innovation in Language Learning and Teaching*, 6(1), 1-13.
- Aki, O. (2006). Is emotional intelligence or mental intelligence more important in language learning? *Journal of Applied Science*, 7(6), 66-70.
- Allan, D. (1992). *Oxford placement test*. Oxford: Oxford University Press.
- Beauvois, M. H. (1992). Computer-assisted classroom discussion in the foreign language classroom: Conversation in slow motion. *Foreign Language Annals*, 25, 455-464.



- Dewaele, J. M., Petrides, K. V., & Furnham, A. (2008). Effects of trait emotional intelligence and sociobiographical variables on communicative anxiety and foreign language anxiety among multilinguals: A review and empirical investigation. *Language Learning*, 59, 911-960.
- Downey, L. A., Mountstephen, J., Lloyd, J., Hansen, K., & Stough, C. (2008). Emotional intelligence and scholastic achievement in Australian adolescents. *Australian Journal of Psychology*, 60, 10-17.
- Ellis, R., & Yuan, F. (2004). The effects of planning on fluency, complexity, and accuracy in second language narrative writing. *Studies in Second Language Acquisition*, 26, 59-84.
- Elola, I., & Oskoz, A. (2010). Collaborative writing: Fostering foreign language and writing conventions development. *Language Learning & Technology*, 14(3), 51-71.
- Foster, P., Tonkyn, A., & Wigglesworth, G. (2000). Measuring spoken language. *Applied Linguistics*, 21(3), 354-75.
- Gardner, H. (1993). Intelligence and intelligences: Universal principles and individual differences. *Archives de Psychologie*, 61, 169-172.
- Goldsworthy, R. (2002). Supporting the development of emotional intelligence through technology. *Computers in the Schools*, 19(1), 119-148.
- Goleman, D. (1995). *Emotional intelligence: Why it can matter more than IQ*. Bantam Books, New York.
- Holsbrink-Engels, G. A. (1997). Computer-based role-playing for interpersonal skills training. *Simulation-and-Gaming*, 28(2), 164-80.
- Housen, A., & Kuiken, F. (2009). Complexity, accuracy and fluency in second language acquisition. *Applied Linguistics*, 30(4), 461-473.
- Hunt, N., & Evans, D. (2004). Predicting traumatic stress using emotional intelligence. *Behavior Research and Therapy*, 42(7), 791-798.
- Kern, R. (1995). Restructuring classroom interaction with networked computers: Effects on quantity and characteristics of language production. *Modern Language Journal*, 79, 457-476.
- Levy, M. (1997). *CALL: context and conceptualization*. Oxford: Oxford University Press.
- Liang, M. Y. (2010). Using synchronous online peer response groups in EFL writing: Revision-related discourse. *Language Learning & Technology*, 14(1), 45-64.
- Lynn, A. B. (2002). *The emotional intelligence activity book. 50 activities for promoting EQ at work*. New York: Amacom.
- MacIntyre, P. D., & Gardner, R. C. (1994). The subtle effects of language anxiety on cognitive processing in the second language. *Language Learning*, 44, 283-305.
- Mayor, J., Salovey, P., & Caruso, D. R. (2004). Emotional intelligence: Theory, findings and implications. *Psychological Inquiry*, 15(3), 197-215.
- Norris, J. M., & Ortega, L. (2009). Towards an organic approach to investigating CAF in instructed SLA: the case of complexity. *Applied Linguistics*, 30(4), 555-78.



- Parker, J. D. A., Summerfeldt, L. J., Hogan, M. J. & Majeski, S. A. (2004). Emotional intelligence and academic success: Examining the transition from high school to university. *Personality and Individual Differences*, 36, 163-172.
- Petrides, K.V., Sangareau, Y., Furnham, A., & Frederickson, N. (2006). Trait emotional intelligence and children's peer relations at school. *Social Development*, 15(3), 537-47.
- Richardson, T. L. (2002). The importance of emotional intelligence during transition into middle school. *Middle School Journal*, 33(3), 55-58.
- Rossiter, M. J. (2003). The effects of affective strategy training in the ESL classroom. *TESL-EJ*, 7(2), 1-20.
- Salovey, P., & Mayer, J. (1990). Emotional intelligence. *Imagination, Cognition, and Personality*, 9(3), 185-211.
- Schutte, S., Malouff, J. M., Bobik, C., Coston, T. D., Greeson, C., Jedlicka, C., et al. (2001). Emotional intelligence and interpersonal relations. *Journal of Social Psychology*, 141, 523-536.
- Smith, B. (2005). English language learning and technology: Lectures on applied linguistics in the age of information and communication technology. *Modern Language Journal*, 89(4), 647-648.
- Song, L. J., Huang, G. H., Peng, K. Z., Law, K. S., Wong, C. S., & Chen, Z. J. (2010). The differential effects of general mental ability and emotional intelligence on academic performance and social interactions. *Intelligence*, 38, 137-143.
- Stottlemeyer, B. G. (2002). *A conceptual framework for emotional intelligence in education: Factors affecting student achievement*. Unpublished doctoral dissertation. Texas: A & M University-Kingsville.
- Yuan, F., & Ellis, R. (2003). The effect of pre-task planning and online planning on fluency, complexity, and accuracy in L2 oral production. *Applied Linguistics*, 24, 1-27.

## Bio Data

**Parisa Abdolrezapour** is a Ph.D. candidate in applied linguistics at the University of Isfahan, Iran, where she also teaches undergraduate courses. She has a number of articles in academic journals like *Discourse Studies*, *Australian journal of linguistics*, *International journal of applied linguistics*, *Innovation in language learning and teaching*, and *Sociolinguistic studies*. Her research interests lie primarily in cognitive aspects of language teaching and learning as well as cross-cultural studies.



## Do Different Stakeholders' Actions Transform or Perpetuate Deleterious High-Stakes Testing Impacts in Iran?

Azizullah Mirzaei<sup>1</sup>

Mahmood Hashemian

Nooshin Tanbakooei

*Shahrekord University*

### Abstract

Messick's conceptualization of consequential validity and the advent of critical language testing have raised a critical awareness among L2 practitioners to scrutinize the educational, social impacts of high-stakes testing practice on tests takers' future careers and lives, on the one hand, and the stakeholders' actions that perpetuate or transform these impacts and the testing status quo, on the other. This critical impact study set out to explore the transformative or reproductive practices different stakeholders (e.g., EFL teachers, learners, and parents) carry out in relation to the detrimental impacts of the Nation-wide University Entrance Test (NUET) in Iran in and out of EFL classrooms. The participants were 60 EFL teachers, a sample of 319 pre-university students, and 15 parents in Isfahan, Shahrekord, and Tehran. A triangulation approach, using multi-faceted questionnaires for EFL students and teachers, a classroom observation checklist, and semi-structured interviews, was adopted to collect the data. The interpretation of the quantitative and interpretative analyses of the data attested to the reproductive nature of the stakeholders' NUET-related practices in Iran. That is, EFL teachers' and (last year) high school students' actions and practices mostly reproduced the negative impacts of the NUET by teaching and learning to the test, practicing test-taking techniques, not using authentic materials, reviewing the previous NUET items, and using them in final/midterm test and quizzes. Finally, it is suggested that NUET stakeholders have to redefine their pedagogical practices in a way that will be able to transform the high-stakes testing situation and the related hazardous test impacts and results in fairer learning-oriented language assessment.

**Keywords:** Consequential validity, Critical language testing, Impacts of high-stakes tests, Stakeholders, Transformative or reproductive practices

### 1. Introduction

At the turn of century there has been an uprising interest to adopt a critical scrutiny of testing and measuring in all aspects. It has been argued that psychometric principles were the core aims of

---

<sup>1</sup> Email: mirzaei-a@lit.sku.ac.ir



traditional testing (Bachman & Palmer, 1996; Weir, 2005). In other words, in the traditional psychometric view, reliability and validity are the major criteria for high quality tests. Nowadays, there has been a shift of trend from traditional testing which guarantees obtained test scores which are valid assessments of learners' abilities and knowledge to critical theories of validity that require test developers to investigate the consequences that tests have on education and societies (Hamp-Lyons, 1997; McNamara, 2000; Messick, 1996; Shohamy, 1998).

This shift of focus from traditional testing to employing more social-ethical views came into existence with the introduction of consequential validity (Messick, 1989) as part of a unified concept of construct validity and critical language testing (CLT) (Shohamy, 2001). Messick (1989) stressed the central importance of the interpretation of the test scores and the social consequences of the actual use of the test scores. Similarly, Shohamy (2000) argues that it is the responsibility of second language testers to problematize not only the tests themselves but also their uses and consequences. One of such consequences that are closely related to consequential validity is washback or backwash phenomenon. The term washback or backwash concerns the effect of tests on teaching and learning as well as on different stakeholders' practices in different situations (Alderson & Wall, 1993; Cheng, 1999; Hamp-Lyons, 1997; Robb & Ercanbrack, 1999). Accordingly, the term itself is neutral in that the influence of a test may be either positive or negative in nature. Some of the researchers in the field have emphasized the negative and detrimental impacts the tests might bring about (Bailey, 1999; Broadfoot, 2005; Cheng, 1999; Madaus, 1988; Shohamy, 1998; Smith, 1991a, 1991b), focusing on the relationship between the content of a test and the resulting narrowing of the curriculum and teachers' training practices. A further concern to the consequential validity and washback phenomenon is the issue of high-stakes tests. Generally speaking, high-stakes tests are those based on which one can make prominent decisions that have educational, financial, or social impacts (Genesee & Upshur, 1996). Considering the vital roles played by the high-stakes tests in societies, recent empirical studies in language education have revealed that high-stakes bring about unfavorable effects on different stakeholders inside and outside of the classrooms (Cheng, 1998, 1999; Luxia, 2005; Shohamy, 1998; Wall, 2000).

The NUET as one of the high-stakes tests in Iran has effects on the test takers' future lives. It has been found that a great number of testing practices and assessments inside and outside the classrooms are centralized on the skills and knowledge that students need to pass this wide scale test. Whatever teachers and students do in classes is heavily under the influence of the content of these tests; that is, whatever is done in high school English classes is summarized towards the reproduction of the negative washbacks of the NUET (that is the successful performance on this test) rather than transformative practices of successful language learning in terms of being able to communicate efficiently. For instance, in English classes in high-schools, speaking and listening



skills are paid but a lip service because they are never tested in the NUET, whereas grammar, vocabularies, and reading comprehension are over emphasized because they are the main parts. Thus, one of the most fruitful areas of research is to scrutinize the negative consequences of the high-stakes NUET on the teaching and learning practices of different stakeholders inside and outside the L2 classrooms.

## 2. Theoretical Background

As proposed by McNamara (2000), testing is one of the universal features of everyone's social life. Whether being aware of it or not, every day people take test in practically every cognitive effort they make. Test is one of the measurement instruments designed to collect information about a specific sample of an individual's behavior (Bachman, 1990; Baker, 1989; Carroll, 1968).

Based on traditional views to language testing, tests are developed based on psychometric principles; that is, the focus was on the accuracy of the tests as following restricted procedures, boundaries, and criteria in order to ensure tests' reliability (i.e., how accurate test scores are) and validity (i.e. the extent to which tests measure what they are expected to measure) (Ainy, 2000; Bachman & Palmer, 1996; Cunningham, 1998; McNamara, 2000). This procedure in turn enables testers and teachers to assemble precise information to make fair and sound professional judgments about test takers. In fact, traditional approaches took an individualized focus in relation to language testing and assessment. Test designers do not consider test takers but in occasions such as computing test indices (e.g., item difficulty, item facility, and item discrimination). According to Shohamy (2001), test takers are classified as to good and bad based on their performances on the tests. She added that "The general rule is that the test takers need to match their performances to the tests rather than the tests to the test takers." (p. 4) Traditional testing is not interested in the issues as how test takers prepare themselves for the tests, how the results of the tests affect test takers' future lives, and what the feelings of test takers are. Shohamy (2001) explained that the traditional test designing is a single and distinct process which is detached from people, society, impacts, and uses.

However, at the turn of the century, there has been an increasing critical awareness in the field of language testing with the aim of shifting attention to the uses of tests in education, society and in the political realm (Andrews, Fullilove, & Wong, 2002; Broadfoot, 2005; Chapman & Snyder, 2000; Hamp-Lyons, 1997; Lynch, 1997; McNamara, 2000; Messick, 1996; Shohamy, 1997). Despite the traditional perspective to language testing, developing a test is not a distinct activity that is developed by a single person who is curious about other individuals, but rather as acts which are embedded in and connected to educational, pedagogical, bureaucratic, psychological, social and



political variables that affect people, knowledge, curriculum, teaching, learning, ethicality, social classes, bureaucracy, politics, inclusion and exclusion (Messick, 1981, 1994, 1996).

Shohamy's (2000) CLT is the result of above mentioned shift which is concerned with what happens to test takers, to the knowledge that is created through tests, and to the teachers who prepare students for tests as well as to the motivations of different stakeholders (e.g., school administrators, EFL teachers, and students) who introduce, administer, and take tests. The CLT poses questions about the effects and impacts of tests on the teaching materials taught and learned as a result of tests and the teaching methods applied in teaching as a reaction to testing. It is concerned with the consequences of test results for parents whose children are subject to tests, ethical and fairness issues associated with the act of testing, and the short and long term consequences of tests to education and societies.

Along with mentioned new theoretical perspectives to language testing, Messick's (1981, 1994, & 1996) notion of consequential validity comprises the rationales and evidences for evaluating the political and social consequences of score interpretations and tests use. These consequences are associated with bias in scoring and interpreting of test scores, with unethical and unfair issues (Kunnan, 2000; Shohamy, 2001; Spolsky 1995), and with positive and negative washback effects and impacts on language teaching and learning (Alderson & Wall, 1993; Bailey, 1999; Messick, 1996; Popham, 1987; Shohamy, Donitsa-Schmid, & Ferman, 1996; Wall, 1997). These mentioned views stand in absolute contrast to the traditional views of language testing.

A further concern in language testing is the relevance of consequential validity and washback (or backwash) and impact phenomena. Washback is sometimes defined differently and, in some extent, contradictory by scholars. All of them agree that washback is the relationship between testing, teaching, and learning, though (Cheng, Watanabe, & Curtis, 2004; Shohamy et al., 1996). Messick (1996) explains that washback is "the extent to which the introduction and use of a test influences language teachers and learners to do things they would not otherwise do that promote or inhibit language learning" (p. 241). Measurement-driven (Popham, 1987), curriculum alignment (Shohamy, 1993), and systematic validity (Fredericksen & Collins, 1989) are the other terms that have similar meanings to washback.

Integral to the concept of washback and impact is the standardized and high-stakes tests. Genesee and Upshur (1996) described high-stakes tests as "the use of tests and assessments alone to make decisions that are of prominent educational, financial, or social impact" (p. 6). Qi (2005) also introduced high-stakes tests as tests whose results are seen by students, teachers, administrators, parents, or the general public as the basis upon which vital decisions that immediately and directly



affect the students' future lives are made. In the same vein, Shohamy et al. (1996) defined a high-stakes test as one used to make decisions about admissions, promotion, placement, or graduation based on test scores while low-stakes exams do not entail these significant decisions. Generally speaking, public examinations or large-scale standardized tests are usually considered as high-stakes tests designed to provide fair and equal chance for all test takers by making the test content, administration condition, and scoring uniform for all test takers, and they are assumed as having a gatekeeper function to access to higher education. Therefore, washback and impacts of standardized, high-stakes tests on stakeholders' teaching and learning behaviors are significantly substantial and controversial. Studies carried out by Alderson and Wall (1993), Wall and Alderson (1993), Alderson and Hamp-Lyons (1996), Shohamy et al. (1996), Bailey (1999), Watanabe (1996a, 1996b), Cheng (1998, 1999), Chapman and Snyder (2000), Andrews et al. (2002), Qi (2005), and Green (2007) which have explored the impacts of high-stakes assessments have documented the negative consequences and washbacks of these tests on teaching, learning, attitudes of different stakeholders, and narrowing the curriculum.

In the last decade, likewise, in the context of Iran enough evidence has been accumulated to convince the researchers of the existence of test impacts on various aspects of education. A brief survey of literature reveals that high-stakes test washbacks on teaching and learning are abundant. Ghorbani (2008) categorized eight factors that highly correlated with the NUET washback effects in Iran: (a) teaching experience, (b) educational background, (c) perceived professionalism in teaching, (d) gender, (e) school type, (f) school location, (g) perceived learners attitudes, and (h) perceived external pressure. He scrutinized the washback effect of the NUET on 377 pre-university English teachers' teaching and curricular planning in six dimensions: (a) syllabus design, (b) methods of teaching, (c) teaching contents, (d) classroom activities and time arrangement, (e) instructional materials, and (f) classroom assessment through questionnaires and interviews. The results indicated that time arrangement was the most affected one because pre-university teachers devoted considerable amount of classroom activities and instruction time to specific test preparation. Moreover, the interviewed teachers acknowledged the negative effect of the NUET and expected the authorities to modify it based on the recent teaching and testing theories. Similarly, Mirzaei and Roshani (2011) applied a triangular approach to investigate the impact of high-stakes testing practices on teaching and learning process of 32 EFL teachers and 70 pre-university students using questionnaires and interviews. The results indicated that the NUET has negative consequences on participants' teaching and learning process in classrooms as teachers in sample employed individual activities, reading aloud technique, and no use of audio-visual aids while pre-university students preferred learning English interactively. In the same vein, Riazi and Razavipour (2011) exploring the impact of centralized English tests on 20 high-school EFL teachers utilizing assessment literacy test and conducting face to face interviews showed that the level of



participants' assessment literacy was lower than expected. Moreover, their test practices were heavily influenced by the content and format of centralized tests, and they prepared learners with less emphasis on speaking and listening skills. A brief survey of literature in the context of Iran revealed the fact that the NUET as one of the significant high-stakes tests has several negative consequential washbacks on different dimensions of education. According to Pishghadam and Mirzaei (2008), teachers are not able to change the rules and teach the materials in books which are prescribed by the Ministry of Education. In other words, teachers as "passive technicians" (Kumaravadivelu, 2003) follow exactly the syllabus and focus on reading, vocabularies, and grammatical rules and neglect the other skills, critical thinking, and learning. Simply put, learners are motivated to memorize decontextualized vocabularies and grammatical rules. In the educational system of Iran, teachers have to transfer the materials presented in textbooks to students. As Hashemi, Naderi, Shariatmadari, Seif Naraghi, and Mehrabi (2010) put it in the educational system of Iran, learners are considered to be like computers, which should be filled with prescribed information that are presented in textbooks.

Accordingly, negative consequences of high-stakes tests have raised questions about the high prices of this testing regime on teaching and learning behaviors and practices of different stakeholders inside and outside the classrooms, which result in perpetuating the negative consequences and washback effects of the NUET, appreciating rote learning, and devaluating critical thinking and creativity. The concept of washback presupposes a belief in the notion that tests are prominent determiners of classroom practices and events (Robb & Ercanbrack, 1999). Unfortunately, the empirical research insights and critiques raised by Shohamy (2001), Spolsky (1995), Lynch (1997) and others are not being sufficiently channeled into testing practices by policy makers, test developers, school administrators, and teachers within the classrooms to change teaching and learning practices and generate more egalitarian alternative testing practices. The power of the current testing establishment is huge. Therefore, there have been many "dark alleys" in which poorly assessment technologies have been used thoughtlessly or have been located on "blind bends" in which a lack of sufficient preparation and knowledge has resulted in assessment practices that are dangerous and, potentially, very damaging to learners (Broadfoot, 2005, p. 123). Broadfoot (2005) argues that there is a need to question and challenge the practicality and desirability of the present situation in a fundamental way and try to examine its consequences for individuals and societies. What testing practice needs in part is a movement toward a transformative model of instruction with a core of critical educators and "transformative teachers" (Kumaravadivelu, 2003) who are willing to change their perspective on certain facts, beliefs, and assumptions (Coombesha & Danaher, 2006) and are employing new theories of testing and ethical and fair tests.



This study sought to explore transformativeness of stakeholders' practices in relation to the high-stakes testing practice inside and outside the classrooms. The researchers adopted a triangulation data collection approach utilizing questionnaires, classroom observations, and interviews. In brief, this study explored the following question through the quantitative analysis of the questionnaire-based data as well as the complementary interpretive analysis of the transcribed interview data.

1. Do different stakeholders (EFL pre-university teachers, pre-university students, and parents) have transformative or reproductive practices in relation to the (probable) negative washback or impact effects of the high-stake tests in Iran?

### **3. Method**

#### *3.1. Participants*

The participants of the study included a random sample of 60 Iranian EFL pre-university teachers, 7 male and 53 female. They were high-school teachers with the age ranged from 23 to 49. In terms of their university degree, 44 held B.A. and 16 others held M.A. Of the EFL teachers participated in this study, five took part in the piloting phase of the teacher questionnaire; additionally, 10 EFL teachers randomly participated in the semi-structured interviews. In addition, a cohort of 319 Iranian math, natural science, and humanities pre-university students from both gender and aged between 17 to 19 participated in this study as the second group. Of the pre-university students in the sample, five took part in the piloting phase of the student questionnaire and 10 were asked in the semi-structured interviews. Finally, 15 Iranian parents were selected in order to participate in the semi-structured interviews. They were asked to participate based on their willingness to take part in the study. It is worth mentioning that gender was not a controlled factor in choosing the participants. Also, the EFL teachers and pre-university students were selected from 18 different high schools in 1<sup>st</sup>, 2<sup>nd</sup>, 3<sup>rd</sup>, 4<sup>th</sup>, and 5<sup>th</sup> city districts in Isfahan, 1<sup>st</sup> and 2<sup>nd</sup> city districts in Shahrekord, and 5<sup>th</sup> and 14<sup>th</sup> city districts in Tehran.

#### *3.2. Instrumentation*

Three types of instruments were employed in this study:

##### *3.2.1. Teacher questionnaire and student questionnaire*

Each Teacher's questionnaire (TQ) and student's questionnaire (SQ) has 18 items and consists of two sections (i.e., reproductive and transformative practices). Also, a four-point Likert scale ranging from Never (1) to Always (4) was used. They were written based on the related literature and



examined by two TEFL experts from Shahrekord University. The questionnaires were piloted in order to obtain information regarding the relevance and clarity of the questions, format, and time allotment. They were then revised accordingly. Moreover, the items in the questionnaires were originally designed in English by the present researchers. Due to the language level of the students in the sample, however, SQ was issued to the pre-university students in Persian to minimize misunderstanding and the reliability and validity problems caused by the language factors (Geisinger, 1994). To enhance the validity of the questionnaires, the Persian version of the questionnaire was back translated into English to see whether anything was missing or could be misinterpreted. For assuring of construct validity, the questionnaires were subjected to principal components analysis (PCA) using statistical package for social sciences software (SPSS 18). Prior to performing PCA, the suitability of the data for factor analysis was assessed. Inspection of the correlation matrix revealed the presence of many coefficients of .3 and above. The value of Kaiser-Mayer-Oklin Measure of Sampling Adequacy (KMO) and the Bartlett's Test of Sphericity for TQ was 0.63 and SQ was 0.737. Cronbach's alpha coefficient was estimated for both questionnaires which resulted in 0.63 (TQ) and 0.57 (SQ), indicating acceptable internal consistency.

### *3.2.2. Classroom observation checklist*

The observation sessions were carried out to address the practices and activities of the different stakeholders inside the classrooms in relation to the negative consequences of the NUET. There were 15 observation sessions in total-ten in high-schools and five in preparatory classes. The observation scheme was designed based on the related literature, consulting two TEFL professors in Shahrekord University, drafting, and revising of the items. A pool of 30 items was expanded. Using the experts' judgments, the researchers modified the original version: 10 items were dropped out on the grounds of being practical, the wording of some was modified, and 2 items were reworded to prevent the Halo effect and to enhance the validity of responses. The checklist was covered of 20 statements in two parts: (a) transformative practice and (b) reproductive practice. Kappa measure of agreement ran to ensure inter-rater reliability and to ascertain the rater consistency in completing the checklists. The Kappa measure of agreement value is .73 ( $p < 0.05$ ). Kappa values above 0.7, according to Peat (2001), represent a good agreement hence a good estimate of inter-rater consistency.

### *3.2.3. Semi-structured interviews*

Finally, based on the related literature, semi-structured interviews consisting of two questions were prepared to ask EFL teachers, pre-university students, and parents. This was done to make sure that the participants could voice their views and opinions unrestricted. The EFL pre-university



teachers and students were interviewed in their respective schools, but parents were interviewed on the date of the NUET (29th, 30th, and 31st of June, 2012).

### *3.3. Data collection procedure*

At first, 10 English female high school classes and 5 English female preparatory classes were observed by one of the researchers in Isfahan, Shahrekord, and Tehran. During each observation, the checklist was completed. The questionnaires were completed by teachers and students in the sample during the class time, which took 15 to 20 minutes, and was preceded by a brief introduction assuring the participants of confidentiality. Face to face interviews focused on the participants' transformative or reproductive practices and activities in relation to negative washback effects and test influences of the NUET inside and outside the classrooms. Each individual interview lasted between 10 to 15 minutes. All the interviews were conducted in Persian and audio recorded. The recordings were transcribed and translated into English by the present authors. The questions asked in the interview centered on the main issue of the study.

## **4. Results**

The present study aimed at investigating the transformative or reproductive practices of different stakeholders (i.e., EFL pre-university teachers, pre-university students, and parents) due to the negative washbacks or impacts of the NUET. The analyses of collected data from TQ, SQ, classroom observation checklists, and semi-structure interviews will be reported and presented in Tables 1 to 10.

### *4.1. Descriptive analysis of TQ and SQ*

First, descriptive statistics were computed for each subcategory of the two questionnaires used in this study. The initial descriptive statistics of the TQ and SQ results (Table 1) reveal that the NUET has considerable negative effects on EFL teachers and pre-university students' teaching and learning behaviors inside and outside of the classrooms, meaning that EFL teachers ( $M = 3.07$ ) and pre-university students ( $M = 2.82$ ) perpetuated the negative aspects of the NUET.

**Table 1.**

Descriptive statistics of the TQ and SQ instruments

Participants	Instruments	Minimum	Maximum	Mean	SD	Kurtosis
	Reproductive practice	1.5	4	3.07	.318	0.25



EFL teachers	TQ	Transformative practice	1	2.87	1.63	.272	1.48
Pre-university students	SQ	Reproductive practice	1	4	2.82	.312	1.18
		Transformative practice	1	2.87	1.20	.214	0.86

#### *4.2. The transformative or reproductive practices of different stakeholders*

##### *4.2.1. The paired-sample t-test of reproductive and transformative practices in TQ & SQ*

Two successive paired-sample *t*-tests were run comparing reproductive and transformative practices in each questionnaire in which the participants determined whether they employed transformative or reproductive teaching and learning practices. These *t*-tests were run in an attempt to test the second hypothesis of the study.

**Table 2.**

Paired-samples t-test of transformative and reproductive practices in TQ and SQ

		<i>M</i>	<i>SD</i>	Mean Difference	<i>SD</i>	<i>t</i>	<i>df</i>	<i>Sig.</i> (2-tailed)
Pair 1	TQ	Reproductive practice	24.61	2.56				
		Transformative practice	13.08	2.18	11.52	3.883	22.79	59
	SQ	Reproductive practice	22.61	2.50	12.978	3.32	69.66	318
		Transformative practice	9.63	1.71				.000

The paired-sample *t*-tests results of the teaching and learning practices of the EFL teachers and pre-university students exhibit a statistically significant difference between transformative and reproductive practices in and out of classrooms. As Table 2 demonstrates, EFL pre-university teachers ( $M = 24.61$ ,  $SD = 2.18$ ) and students ( $M = 22.61$ ,  $SD = 2.50$ ) favored reproductive practices in their English classrooms when the negative washbacks of the NUET are considered. The eta squared statistic for TQ was 0.60 and for SQ was 0.93 indicated large effective sizes.



#### 4.2.2. The chi-square results of reproductive and transformative practices in TQ and SQ

With the intention of having a closer analysis of transformative and reproductive sections of the TQ and SQ, four Chi-square tests for independence were run. The Chi-square results are illustrated in Table 3-6.

**Table 3.**

EFL pre-university teachers' crosstabulation of reproductive practices

Sections	Items & Dimensions	Items 1-8 TQ				Chi-square	df	Asymp. Sig.
		Never 1	Sometimes 2	Often 3	Always 4			
Reproductive practices	1. Teaching the contents of the NUET	0%	6.7%	23.3%	70%	17.15	2	.000
	2. Suggesting students to prepare the helpful complementary materials	5%	25%	48.3%	21.7%	14.01	3	.003
	3. Being judged by the results of the NUET	5%	16.7%	35%	43.3%	2.50	3	.475
	4. Assessing students with multiple-choice grammatical and vocabulary items	3.3%	3.3%	38.3%	55%	165.9	3	.000
	5. Bringing relevant materials to the class	0%	0%	16.7%	83.3%	16.80	1	.000
	6. Teacher's responsibility to help students pass the NUET	8.3%	10%	53.3%	28.3%	16.09	3	.001



	7. Reviewing the previous NUET items	0%	1.7%	50%	83.3%	37.87	2	.000
	8. Suggesting students to participate in preparatory classes	5%	26.7%	56.7%	11.7%	29.09	3	.000

**Table 4.**

EFL pre-university teachers' crosstabulation of transformative practices

Sections	Items & Dimensions	Items 8-16 TQ				Chi-square	df	Asymp. Sig.
		Never 1	Sometimes 2	Often 3	Always 4			
Reproductive practices	9. Participating in a community of teachers and school administrators to talk about the NUET critically	83.3%	11.7%	5%	0%	192.7	2	.000
	10. Participating in a community of teachers and school administrators to talk about the students' educational and social problems result from the NUET	44%	51%	5%	0%	58.74	2	.000



	11. Talking to parents about the students' affective and motivational problems result from the NUET	35%	51.7%	11.7%	1.7%	96.34	3	.000
	12. Writing articles about the negative impacts of the NUET to EFL/ESL conferences	100%	0%	0%	0%			
	13. Doing research on inequality and ethical problems ensue from the NUET	98.3%	1.7%	0%	0%	114	1	.000
	14. Writing letters to the Ministry of Education about the negative impacts of the NUET	95%	3.3%	1.7%	0%	319.3	2	.000
	15. Changing the direction of teaching English to increase students' critical thinking	11.7%	58.3%	20%	10%	59.75	3	.000
	16. Considering students from low-income families who cannot buy commercial books	66.7%	33.3%	11.7%	1.7%	30	1	.000

**Table 5**

Pre-university students' crosstabulation of reproductive practices

Sections	Items & Dimensions	Items 1-8 from SQ				Chi-square	df	Asymp. Sig.
		Never 1	Sometimes 2	Often 3	Always 4			
Reproductive practices	1. Learning the contents of the NUET	0.9%	3.4%	36.7%	58.9%	103.2	3	.000
	2. Buying the helpful complementary materials	6.9%	29.5%	50.5%	13.2%	119.3	3	.000
	3. Judging the teacher by the result of the NUET	2.2%	9.1%	50.5%	38.2%	83.22	3	.000
	4. Being assessed with multiple-choice grammatical and vocabulary items	0.3%	4.7%	31%	63.9%	1141	3	.000
	5. Reviewing relevant materials to the NUET in the class	3.8%	19.1%	58.6%	18.5%	136.5	3	.000
	6. Teacher's responsibility to help us pass the NUET	3.1%	10.3%	44.8%	41.7%	119.6	3	.000
	7. Reviewing the previous NUET items	1.6%	6%	31.7%	60.8%	88.98	3	.000
	8. Participating in preparatory classes	10.7%	14.4%	45.8%	29.2%	92.12	3	.000



**Table 6**

Pre-university students' crosstabulation of transformative practices

Sections	Items & Dimensions	Items 8-16 from SQ				Chi-square	df	Asymp. Sig.
		Never 1	Sometimes 2	Often 3	Always 4			
Transformative practices	9. Participating in a community of teachers and school administrators to talk about the NUET critically	89%	6.9%	3.1%	0.9%	2218	3	.000
	10. Participating in a community of teachers and school administrators to talk about our educational and social problems result from the NUET	87.1%	11.3%	1.6%	0%	1146	2	.000
	11. Talking to parents about our affective and motivational problems result from the NUET	76.5%	21%	0.9%	1.6%	1617	3	.000
	12. Teacher wrote articles about the impacts of the NUET to EFL/ESL conferences	99.7%	0.3%	0%	0%	632	1	.000



Sections	Items & Dimensions	Items 8-16 from SQ				Chi-square	df	Asymp. Sig.
		Never 1	Sometimes 2	Often 3	Always 4			
	13. Teacher did research on inequality and ethical problems ensue from the NUET	100%	0%	0%	0%			
	14. Writing letters to the Ministry of Education about the negative impacts of the NUET	100%	0%	0%	0%			
	15. Changing the direction of learning English to increase our critical thinking	49.8%	44.2%	4.1%	1.9%	787.1	3	.000
	16. Teacher considered students from low-income families who cannot buy commercial books	61.1%	32%	6%	0.9%	1039	3	.000

Table 3 demonstrated that most of the EFL pre-university answered reproductive practice section with “Often” or “Always”, meaning that the participants reproduce the negative washback of the high-stake test like the NUET. Teachers in the sample opted for “Never” or “Sometimes” in transformative practice part (Table 4). That is, they do not change their teaching practices inside the classrooms.



As it is seen in Table 5, most of the pre-university students completed SQ with “Always” or “Often” in *reproductive practice* part. A closer examination of the results shows that most of the pre-university students in the sample used “Never” or “Sometimes” for items in *transformative practice* part (Table 6). In fact, a considerable number of participants depicted reproductive practices inside and outside the classroom as they perpetuate the negative washbacks of the NUET.

In sum, based on the obtained results, it can be safely assumed that the majority of the EFL pre-university teachers and pre-university students who participated in the study do not try to shift their focus from linguistic knowledge to the teaching and learning of language use. These stakeholders reproduce the negative washback effects of the NUET as one of the high-stake tests in Iran by teaching and learning different sections of the test. Teaching to the test content and test techniques also implies that English Language learning is a skill of grammar and vocabularies that should be practiced and learned, which results in unwanted washback, as illustrated by the case of the NUET. Thus, the null hypotheses of this study were rejected.

- H0<sub>1</sub>: Different stakeholders’ educational practices neither transform nor reproduce the (probable) negative washback or impacts of the NUET in Iran.

#### *4.2.3. The results of transformative practices and reproductive practices in classroom observation checklists*

The successive of *transformative practice* and *Reproductive practice* of the classroom observation checklists are used as approval documents for the previous outcomes. A closer examination of the results shows that the EFL pre-university teachers and students in the samples try to adjust their teaching and learning practices and activities to the requirement of the NUET. The evidence to such a claim is teaching test-taking strategies (i.e., test-wiseness and test tricks), ignoring the cultural aspect of language learning and neglecting of idiomatic expressions, and encouraging their students to use the commercial textbooks (e.g., Gaj, Kanun-e Farhangi-e Amuzesh, and Katab-e Sefid). The greatest concern of the teachers was the pre-university students’ rote memorization of synonyms, the Persian equivalents of the new vocabularies, and explanation of grammatical rules. It could be suggested, then, that the results of the checklists serve to confirm the findings of the TQ and SQ. In other words, to a large extent, EFL pre-university teachers and pre-university students would appear to be doing at classroom level what they, in fact, claim to be doing. Another telling feature of the sets of data within this table is the absence—across all 15 sessions—of speaking and listening skills, cultural aspects of language learning, or activities which facilitate speaking and listening skills.



#### *4.2.4. The semi-structured interview results related to transformative or reproductive practices*

Similarly, the participants' responses to the transformative or reproductive question in the semi-structured interviews further supported the previous findings in TQ, SQ, and classroom observation checklists. The results largely showed that EFL pre-university teachers, students, and parents perpetuate and reproduce the negative washback effects of the NUET.

*Teacher 3 (Female, Isfahan, 1<sup>st</sup> Ds.):*

I definitely spend a great amount of class time on teaching grammatical rules and vocabulary in L1 because I have to do it based on the needs of the pre-university students and lesson syllabuses. If you have a look at the English books in this level, you will find just a huge amount of grammatical rules and vocabularies. There is no room and need for practicing communication in high schools.

*Teacher 6 (Female, Shahrekord, 2<sup>nd</sup> Ds.):*

As the demands of society, school administrations, parents, and especially students themselves, the core aim of language learning in schools is being successful in tests and the NUET. That is why my focus is on grammar and vocabularies and generally tests' contents. Nothing will be changed unless the Ministry of Education changes its policies.

*Student 8 (male, Tehran, 5<sup>th</sup> Ds.):*

Guidance school and high school English classes are the places for practicing grammar and vocabularies. Our teacher forces us to memorize the vocabularies and grammatical rules. So, it is not necessary to practice communication there because we want to pass the NUET, and [laughing] in the NUET you do not need listening or speaking skills. I mean you can learn them in language institutions.

*Parent 3 (male, Tehran):*

I am not a professional person in this field, but, based on my studies, our testing regime need to be changed. Students' single performances on the Konkoor are not reasonable criteria. There should be ongoing assessments during their education. Education in pre-university period is along with the content of Konkoor. Whatever they do or practice is just for being successful in this test.



## 5. Discussion

The current study attempted to examine the transformative or reproductive practices of different stakeholders (i.e., EFL teachers, pre-university students, and parents) in relation to the negative consequences of the NUET. The results of previous studies endorsed the existence of negative washback effects of the NUET on different stakeholders (e.g., school administrators, EFL teachers, pre-university students, and parents), teaching methodology, and curriculum (Ghorbani, 2008; Mirzaei & Roshani, 2011; Riazi & Razavipour, 2011; Roshani, 2008). For example, EFL teachers teach the contents of the NUET, testing strategies, grammar, and vocabularies. Furthermore, they include the NUET in their final and midterm exams. Pre-university students wanted to know as much as possible about the test, format, and contents of the NUET, and they believed it was the teacher's responsibility to provide them this information. Findings of related literature revealed that the focus of teachers and students in high school is how to obtain a high score in the NUET rather than the desire to communicate in English in Iran. Since the NUET functions as a high-stakes gate-keeping test which has serious consequences for the future careers and lives of test takers, its impact on EFL pre-university teachers and pre-university students' teaching and learning behavior is inevitable.

Based on the results of the data analysis, the negative washback of the NUET serves to inform and shape the practices and educational behaviors of the teachers and students. So, most EFL pre-university teachers spent much of the class time on the contents of the NUETs of the test, prepared students in how to pass the test instead of teach how to communicate, reviewed and worked on the supplementary materials and previous NUETs, and ignored the importance of four language skills integration. Furthermore, students are encouraged to work individually rather than cooperatively, memorize grammatical rules and decontextualized vocabulary items that are useful for their future performances, and learn several test strategies or test tricks to answer the multiple-choice items in the NUET. Moreover, the obtained data from the classroom observation checklists and semi-structure interviews revealed that teachers tend to adopt certain tasks and activities that they think are useful for their students' success in future. They also relied on previous NUETs as models for their students to practice the probable content of the tests. Additionally, being able to pass the NUET was regarded more important than being able to communicate in English by most pre-university students. Indeed, they expected their teachers to work on issues that contain the content of the test, they paid attention to speaking and listening skills a lip service because they are not tested in the NUET, and they mostly participated in the preparatory classes.

As the main issue of this study, it was tried to investigate the transformative or reproductive practices of the participants in relation to the negative washback effects of the NUET. Given the



nature of TQ results, then, it can be suggested that EFL pre-university teachers recognize the importance of the NUET and they are more willing to use practical techniques which can aid that development. In sum, the results of this study further support the argument of previous ones. It can be concluded that teachers and students perpetuated the negative washbacks of the NUET by assigning a high weight to the structure of the language, vocabularies, reading comprehensions, and cloze passages. No one has ever tried to question this testing regime by employing transformative pedagogy and holding meeting and conferences about the social, emotional, and educational effects of the test by either teachers or students.

With regards to the nature of the NUET, it is not surprising to see that most teachers and students believed it was very important to practice grammar and vocabularies. It has also been claimed that EFL pre-university teachers in Iran have found themselves in an undesirable position in which the constraints imposed by the examination-driven "hidden syllabus" prevent them from implementing communicative methodology which results in reproducing the negative washbacks of the NUET. In most occasions, however, stakeholders seem less likely to embrace critical teaching/learning, new ideas, or methods simply because of the format of the NUET, preferring instead to continue teaching and learning the grammar, decontextualized grammar with which they are most familiar.

Therefore, it can be argued that EFL teachers, rather than learners, are far more likely to be able to insert the necessary changes and transformative pedagogy in order to improve the actual teaching, learning, and testing process, at least within the context of the classroom. Generally, the core principle of critical pedagogy is "the pedagogy of hope and understanding, since without the possibility of change and a willingness to change criticism does not make much sense" (Akbari, 2008, p. 282). So, there is a crucial demand for conceptualizing testing and measuring because the notion of transformative pedagogy remains incompatible with the traditional educational system in Iran.

In summary, Shohamy (2001) has named tests as instruments of power and control that result in bias, undemocratic, unfair testing practice and boring mechanical teaching and learning activities. Then, high-stakes tests play a central role in students' future educational lives. The higher the stakes, the more stakeholders focus instruction on the tests. Consequently, what is not tested often is not taught. Several aspects of language (e.g., speaking, listening, cultural aspects of language, idioms, implicatures, and speech acts) will be disregarded. Regarding the findings of this study which are in congruent with other related studies (Ghorbani, 2008; Mirzaei & Roshani, 2011; Mohamadi, 2007; Roshani, 2008), there seems little doubt that the NUET has negative washback effects and widely influenced different stakeholders (e.g., EFL pre-university teachers, pre-university students, and school administrators), practices, and educational activities. Additionally, most of the



surveyed and interviewed EFL pre-university teachers, regardless of their experience, educational background, gender, and the school location where they taught, agreed that the NUET has considerable negative consequences on their teaching activities and practices. In most cases, they indicated that they incorporated test taking techniques and strategies into their existing teaching practices rather than employing a communicative approach. Likewise, pre-university students go under heavy memorization of vocabularies, grammatical rules, and correct translation of reading passages at the expense of communicative language learning.

Such studies call for a more critical language testing view toward testing and assessments in the context of Iran educational system. Thus, spending millions of Rials on training English teachers and improving their level of knowledge at teacher training colleges and universities would be a great loss if the present testing practice in Iran remains unchanged. Therefore, it is desirable to consider teaching, learning, and testing reforms based on the latest theories of critical pedagogy and critical language testing to minimize the negative impacts of high-stakes testing practice on societies and students' educational lives. It is hoped that the findings of this study empower EFL teachers and pre-university students, who are in a weak position in an academic hierarchy, to improve the entrance exam system by their viewpoints and to question and challenge the system of teaching and testing practice in Iran. It is tried to encourage them to change their teaching and testing practices and not just reproduce the previous ones.

## References

- Ainy, S. (2000). Some alternatives to the traditional paper and pencil language tests in English. *Journal of NELTA*, 5(2), 30-42.
- Akbari, R. (2008). Transforming lives: Introducing critical pedagogy into ELT classrooms. *ELT Journal*, 62(3), 276-283.
- Alderson, J. C., & Hamp-Lyons, L. (1996). TOEFL preparation courses: A study of washback. *Language Testing*, 13, 280-297.
- Alderson, J. C., & Wall, D. (1993). Does washback exist? *Applied Linguistics*, 14, 115-129.
- Andrews, S., Fullilove, J., & Wong, Y. (2002). Targeting washback—A case study. *System*, 30, 207-223.
- Bachman, L. F. (1990). *Fundamental considerations in language testing*. Oxford: Oxford University Press.
- Bachman, L. F., & Palmer, A. S. (1996). *Language testing in practice*. Oxford: Oxford University Press.
- Bailey, K. M. (1999). *Washback in language testing*. TOEFL Monograph Series, Princeton, NJ: Educational Testing Service.



- Baker, E. L. (1989). *Can we fairly measure the quality of education?* (Tech. Rep. No. 290). Los Angeles: University of California, Center for the Study of Evaluation.
- Broadfoot, P. M. (2005). Dark Alleys and Blind Bends: Testing the language of learning. *Language Testing Journals*, 22(2), 123-141.
- Chapman, D. W., & Snyder, C. W. J (2000). Can high stakes national testing improve instruction: Reexamining conventional wisdom. *International Journal of Educational Development*, 20, 457-474.
- Cheng, L. (1998a). *The washback effect of public examination change on classroom teaching: An impact study of the 1996 Hong Kong certificate of education in English on the classroom teaching of English in Hong Kong secondary schools*. Unpublished doctoral dissertation, University of Hong Kong, Hong Kong.
- Cheng, L. (1998b). Impact of a public English examination change on students' perceptions and attitudes toward their English learning. *Studies in Educational Evaluation*, 24, 279-301.
- Cheng, L. (1999). Changing assessment: Washback on teacher perceptions and actions. *Teaching and Teacher Education*, 15, 253-271.
- Cheng, L., Watanabe, Y., & Curtis, A. (Eds.). (2004). *Washback in language testing: Research contexts and methods*. London: Lawrence Erlbaum.
- Coombes, P. N., & Danaher, G. R. (2006). From the margins to the centre: The power of transformative learning in Australia. *Teaching and Teacher Education*, 22(7), 759-765.
- Cunningham, C. R. (1998). *The toeic test and communicative competence do test score gains correlate with increased competence?* Retrieved from World Wide Web, June 12, 2011, <http://www.birmingham.ac.uk/Documents/collegeartslaw/cels/essays/matefltesdissertations/Cundiss.pdf>
- Geisinger, K. F. (1994). Cross-cultural normative assessment: Translation and adaptation issues influencing the normative interpretation of assessment instruments. *Psychological Assessment*, 6(4), 304-312.
- Genesee, F., & Upshur, J. (1996). *Classroom-based evaluation in second language education*. Cambridge University Press.
- Ghorbani, M. R. (2008). The washback effect of the university entrance examination on Iranian English teachers' curricular planning and instruction. *The Iranian EFL Journal*, 2, 60-87.
- Green, A. (2007). Washback to the learners: Learners and teacher perspectives on IELTS preparation course expectation and outcomes. *Assessing Writing*, 11, 113-134.
- Hamp-Lyons, L. (1997). Washback, impact and validity: Ethical concerns. *Language Testing*, 14(3), 295-303.
- Hashemi, S. A., Naderi, E., Shariatmadari, A., Seif Naraghi, M., & Mehrabi, M. (2010). Science production in Iranian educational system by the use of critical thinking. *International Journal of Instruction*, 3(1), 61-76.



- Kumaravadivelu, L. (2003). *Beyond methods: Macrostrategies for language teaching*. Yale University Press.
- Kunnan, A. J. (2000). *Fairness and validation in language assessment: Selected papers from the 19<sup>th</sup> language testing research from Colloquium, Orlando, Florida Studies in Language Testing*. Cambridge University Press.
- Luxia, Q. (2005). Stakeholders' conflicting aims undermine the washback function of a high-stakes test. *Language Testing*, 22(2), 142-173.
- Lynch, B. (1997). In search of ethical test. *Language Testing*, 14(3), 315-327.
- Madaus, G. F. (1988). The influence of testing on the curriculum. In L. N. Tanner (Ed.), *Critical issues in curriculum: Eighty-seventh yearbook of the National Society for the Study of Education*, (pp. 83-121). University of Chicago Press.
- McNamara, T. (2000). *Language testing*. Oxford University Press.
- Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational measurement* (3<sup>rd</sup> ed.) (pp. 13-103). New York: National Council on Measurement in Education and American Council on Education.
- Messick, S. (1994). The interplay of evidence and consequences in the validation of performance assessments. *Educational Researcher*, 23(2), 13-23.
- Messick, S. (1996). Validity and washback in language testing. *Language Testing*, 13(3), 241-256.
- Mirzaei, A. & Roshani, N. (2011). A critical study of the impact of high-stakes testing practice on the teaching-learning process. *English Language Assessment*, 6, 75-102.
- Mohammadi, M. (2007). High-Stakes testing washback: A survey on the effect of Iranian MA entrance examination on teaching. Retrieved October 10, 2011, from World Wide Web: <http://faculty.ksu.edu.sa/aljarf/documents/english%20language%20teaching%20conference%20-%20iran%202008/mojtaba%20mohammadi.pdf>
- Peat, J. (2001). *Health science research: A handbook of quantitative methods*. Sydney: Allen & Unwin.
- Pishghadam, R., & Mirzaei, A. (2008). English language teaching in postmodern era. *TELLSI*, 2(7), 89-109.
- Popham, W. J. (1987). The merits of measurement-driven instruction. *Phi Delta Kappa*, 68, 679-682.
- Qi, L. (2005). Stakeholders' conflicting aims undermine the washback function of a high-stakes test. *Language testing*, 22, 142-173.
- Riazi, M., & Razavipour, K. (2011). In agency of EFL teachers under the negative backwash effect of centralized tests. *IJLS*, 5(2), 123-142.
- Robb, T., & Ercanbrack, J. (1999). A study of the effect of direct test preparation on the TOEIC scores of Japanese university students. *TESL-EJ*, 3(4). Retrieved on 28 June 2012.



- Roshani, N. (2008). *The impact of the high-stakes test of Konkoor on the congruity of students preferred learning styles with EFL teachers employed teaching styles and the effect of this on students learnacy*. Unpublished Master Thesis, Alzahra University of Tehran.
- Shohamy, E. (1993). The exercise of power and control in the rhetorics of testing. In A. Huhta, K. Sajavaara, & S. Takala (Eds.), *Language testing: New openings*. (pp. 23-38). University of Jyväskylä: Institute for Educational Research.
- Shohamy, E. (1998). Critical language testing and beyond. *Studies in Educational Evaluation*, 24(4), 331-345.
- Shohamy, E. (2000). Fairness in language testing. In A. Kunnan (Ed.), *Fairness and validation in language assessment*. Selected papers from the 19<sup>th</sup> language testing research from Colloquium, Orlando, Florida (pp. 15-19). Cambridge University Press.
- Shohamy, E. (2001). *The power of tests: A critical perspective on the uses of language tests*. London: Longman.
- Shohamy, E., Donitsa-Schimid, S., & Ferman, I. (1996). Test impact revisited: Washback effect over time. *Language Testing*, 13, 298-317.
- Smith, M. L. (1991a). Meanings of test preparation. *American Educational Research Journal*, 28, 521-542.
- Smith, M. L. (1991b). Put to the test: The effects of external testing on teachers. *Educational Researcher*, 20(5), 8-11.
- Spolsky, B. (1995). The examination of classroom backwash cycle: Some historical cases. In D. Nunan, V. Berry, & R. Berry (Eds.), *Bringing about change in language education* (pp. 55–66). University of Hong Kong, Department of Curriculum Studies.
- Wall, D. (1997). Impact and washback in language testing. In C. Clapham & D. Corson (Eds.), *Language testing and assessment* (pp. 291-302). Encyclopedia of language and education, Vol. 7. Dordrecht: Kluwer Academic.
- Wall, D. (2000). The impact of high-stakes testing on teaching and learning: Can this be predicted or controlled? *System*, 28, 499-509.
- Wall, D., & Alderson, J. C. (1993). Examining washback: The Sri Lankan impact study. *Language Testing*, 10, 41-69.
- Watanabe, Y. (1996a). Investigating washback in Japanese EFL classrooms: Problems of methodology. In G. Wigglesworth & C. Elder (Eds.), *The language testing circle: From inception to washback* (pp. 208-239). Melbourne, Victoria, Australia: Applied Linguistics Association of Australia.
- Watanabe, Y. (1996b). Does grammar translation come from the entrance examination? Preliminary findings from classroom-based research. *Language Testing*, 13, 318-333.
- Weir, C. J. (2005). *Language testing and validation: An evidence-based approach*. New York: Palgrave Macmillan.



## Bio Data

**Azizullah Mirzaei** is assistant professor of Applied Linguistics at Shahrekord University, Iran. He has published or presented papers in (inter)national journals (e.g., TELL, IJAL, JTSL, Korean Language Assessment, Educational Psychology) or conferences (TELLSI, ILI, and KELTA). He recently co-authored a chapter on Assessing SL Pragmatics in The Cambridge Guide to SL Assessment. His research interests include: Sociocultural Theory and SLL, L2 Fluency, Interlanguage Pragmatics, Language Testing and Assessment, and Teacher Education.

**Mahmood Hashemian** is assistant professor at Shahrekord University. He has published articles in academic journals such as IJAL, IJLS, JALS, Linguistik Online, JLTR, TPLS, Iranian EFL Journal, and International Journal of Social Sciences. Also, he has given lectures in several national and international conferences such as TELLSI (4, 7, 8, & 9), LDP2010 (Ahvaz, Iran), ELT in the Islamic World (Tehran, Iran), 2<sup>nd</sup> International Language Conference 2011 (Malaysia), the 1<sup>st</sup> Thammasat ELT International Conference: Voices in ELT (Thailand), and 1<sup>st</sup> International AKDENIZ Language Studies Conference 2012 (Turkey).

**Nooshin Tanbakooei** got her B.A. in TEFL from Najaf Abad University and is now an M.A. student in TEFL at Shahrekord University, Iran. She is also teaching English at some language institutes in Isfahan. Her areas of interest include Applied Linguistics, Discourse Analysis, Interlanguage Pragmatics, Language Testing and Assessment.



## Language Teaching and Iranian EFL Learners: Restrictions in Verb Selection and Errors

**Marziye Rabieefar<sup>1</sup>**

*Payam-e-Noor University, Rasht*

**Rajabali Askarzadeh Torghabeh**

*Ferdowsi University of Mashhad*

### Abstract

Iranian students, who learn English as a foreign language, make mistakes and errors in learning the skills of the new language. This paper tries to have a review of concepts and theories concerning error analysis and presents types of errors in which Iranian EFL Learners make. It focuses on error analysis and studies written errors through subcategorization frame in Iranian EFL learners in elementary and intermediate levels. Seventy students from different institutes were asked to translate some sentences into English. The aim of this study is to find out whether lack of L2 subcategorizational information will cause errors in production of L2 and if increasing proficiency level eliminates the errors. Findings of this study is supposed to confirm that lack of L2 subcategorizational information would be a major cause of errors which are highly affected by proficiency level.

**Keywords:** Error analysis, Subcategorization, Iranian EFL learners, Proficiency level

### 1. Introduction

EFL teachers enjoy hearing their students produce flawless English. Unfortunately it is inevitable to avoid errors of different types. Familiarity with the types of errors students make is important in order to prepare instruction and implement it in the EFL classroom. Approving this notion has led emergence of large number of empirical research on adult foreign language errors (Burt, 1975).

According to Brown (1987) errors were considered the result of ineffective learning of the rules of the target language in the behaviorist theory of language learning.

---

<sup>1</sup> Email: rabieefar@gmail.com



In the early 1950's, language was accepted as a system and second language learning as acquisition of two language systems. Errors were regarded as evidence of language transfer, and were seen as the result of ineffective language learning and their elimination became the intension of linguists and language teachers (Khodabandeh, 2007).

In the late 1960's, it was demonstrated that memorizing target language rules and using them to produce language were not all language learners did, they were also constructing their own rules based on the input they had received and that led to a rebirth of error analysis which needed redefinition of the whole concept a more cognitive perspective (Khodabandeh, 2007).

Error analysis was used as a diagnostic tool, helping to identify the causes of errors. Crystal (1980), trying to redefine error, attributed mistakes to a simple pause, metanalysis, or a 'slip of the brain'. Hence, errors became an aid for describing and explaining how language learners acquired a language rather than how they followed a set of rules assumed to be standards of expression.

Today, Khodabandeh (2007) defined error analysis as a kind of linguistic analysis consisting comparison and focusing on errors learners make and believed error analysis has manifested for identifying, classifying and interpreting the language learners' mistakes and has associated with hypotheses and theories of language learning. Richard (1974) held that error analysis at classroom level could be a means of assessing learning and teaching.

Error analysis, as to Brown (1994, p. 204), "emphasizes the significance of errors in learners' interlanguage system". Khodabandeh (2007, p. 8) has cited definition of this term from some scholars:

The term interlanguage, according to Selinker (1972), refers to the systematic knowledge of an L2 which is independent of both the learner's L1 and the target language. Nemser (1971) referred to it as the Approximate System, and Corder (1967) as the Idiosyncratic Dialect or Transitional Competence.

Burt (1975) differentiated 'error analysis approach' from 'contrastive analysis' in that "error analysis does not assume that first language interference is the major predictor of adult errors. Rather, no assumptions are made about the causes of error types" (p. 54). Findings of researches on foreign language errors indicated that interference errors are only one of the types of errors observed in the syntax, morphology and lexicon of target language produced by learners (Ervin-Tripp, 1970; George, 1972; & Grauberg, 1971; Politzer, 1974; Richards, 1971, as cited in Burt, 1975).



For example, Grauberg (1971), found that “mother tongue interference could account for only 25% of the lexical errors, 10% of the syntactic errors, and none of the morphological errors in his students’ essays” (Burt, 1975, p. 54).

Keshavarz (1994) put errors in two main categories; the first category as Syntactical-morphological errors including wrong use of prepositions, articles, plural morphemes, qualifier and intensifier, and the use of typical Persian construction in English. The second category as Lexical-semantic errors including cross association and language switch.

The primary causes of errors reviewed by Khodabandeh (2007, p. 8) are:

Interlingual/Transfer errors: those attributed to the native language (NL). There are interlingual errors when the learner’s L1 habits (patterns, systems or rules) interfere or prevent him/her, to some extent, from acquiring the patterns and rules of the second language (Corder, 1971). Interference (negative transfer) is the negative influence of the mother language (L1) on the performance of the target language learner (L2) (Lado, 1964).

Intralingual/Developmental errors: those due to the language being learned (TL), independent of the native language. According to Richards (1970) they are items produced by the learner which reflect not the structure of the mother tongue, but generalizations based on partial exposure to the target language. The learner, in this case, tries to “derive the rules behind the data to which he/she has been exposed, and may develop hypotheses that correspond neither to the mother tongue nor to the target language” (Richards, 1974, p. 6).

Zobl (1980), however, discussed that origins of both types of errors is processing of properties of L2 input. This division has become the basis of many studies in the domain of errors analysis and transfer studied.

Dulay and Burt (1972) have demonstrated the traditional distinction of developmental vs. transfer errors in Table 1.



**Table1.**

Distinguishing features of developmental and transfer errors

Product Level	Developmental	Transfer
	Errors are identical to those made by native learners (Mougeon & Hebrard, 1975)	Interference yields different types of errors than those that are developmental (Dulay & Burt, 1974a)
	Errors result from the creation of rules similar to those of native learners (Dulay & Burt, 1974a)	Errors reflect learner's use of L1 as a crutch at low level of L2 proficiency (Taylor, 1974)
	Errors reflect learner's competence at a particular developmental stage (Richards, 1971a)	Errors reflect use of L1 as a source of hypotheses about L2 (e.g., Cancino, Rosansky & Schumann, 1975 )
	Errors illustrate learner's attempt to build up hypotheses about the L2 (Richards, 1971a)	
Process Level		
	Origins of errors are within the L2 ( Richards 1971a)	Errors reflect inability to separate the two languages (Richards, 1971a)
	Errors reflect general characteristics of rule learning and language acquisition (Richards, 1971a)	Errors result from L1 habits (out-dated Contrastive Analysis position)
	Main strategies are simplification, generalization and reduction of grammatical redundancy (Dulay & Burt, 1972)	Errors represent an interlingual generalization (Taylor, 1974)

Excerpted from Zobl (1980)

There is also another widely known division of errors, global vs. local errors. Burt (1975) suggested four aspects of English grammar that often cause global errors: basic word order, sentence connectors, psychological predicate constructions and selectional restrictions on certain types of verbs in sentential complements. He believed that when learners need to produce a sentence they had not heard before merely by applying grammatical rule they had already acquired, they would make errors. Errors of this kind clarifies that L1 is not the source, rather it English itself which causes errors.

As proficiency level increases and learners get mastery over correct forms errors reduce and students excel in producing language. Brown (1994) also found that a large number of errors



happen at the beginning levels of language learning, due to negative transfer. He concluded that as language learners improve, intralingual errors overcome interlingual errors.

In this sense, Grains and Redman (as cited in Li & Chan, 1999, p. 85) suggest the term “grammare of vocabulary”. Examples of this include misuse of verb transitivity (\*to discuss about something), confusion between a noun and verb (\*to emphasise on something), preposition (\*to walk under the sun), and so forth.

This grammar of vocabulary is what we are investigating in this paper. In the present study we focus on verb selectional restriction within subcategorization frame. The following questions will be answered:

1. Do subcategorizational differences between Persian and English lead to inaccuracy in production of English as L2 in institutes?
2. In case they lead to errors, do they change to correct form as the proficiency level increases?

## **2. Theoretical Background**

Most of the current studies in error analysis investigate either the linguistic classification of errors or on the causes of errors made by adults learning English and other foreign languages. For instance, Sadeghi (2009) performed a study on collocational differences between L1 and L2 and documented that 72.1% of high school students failed to use collocations correctly. He further traced 83.75% of errors to interlingual interference (Persian as L1) and 16.25% of errors to intra-lingua interference (English or other factors).

Behjat and Sadeghi (2010) investigated if the Threshold Hypothesis can account for Iranian EFL learners' grammar development at different levels and found out that transfer of L1 grammar operates differently at different stages.

Abbasi and Karimian (2011) investigated grammatical errors among Iranian Translation Students according to Keshavarz's (1994) model. Their findings surprisingly showed that 98 percent of the students had grammatical problems. Furthermore, most of errors were of interlingual errors, indicating the influence of the mother language.

In a study conducted by Nayernia (2011), written sentences of learners were analyzed to find out what proportion of the learners' errors were intralingual errors and whether the native language



plays a significant role in learners' difficulties in learning the target language. Her findings revealed that only 16.7 percent of the errors were interlingual errors and most of errors could be attributed to target language system. This is in contrast with Abbasi and Karimian's (2011) finding.

Namvar et al. (2012) analyzed collocations in the Iranian postgraduate students' writings to explore the influence of first language (L1) and the cultural background of learners on the production of collocations. Writing and found out that first language influence appeared to have a strong effect on the learners' production of collocation.

In another study Sattari (2012) analyzed Persian English learners' grammatical errors in writing and documented that learners at elementary levels made a great number persistent errors which could be traced to the mother tongue.

### **3. Method**

#### *3.1. Participants*

A group of seventy learners, forty at elementary level and thirty at intermediate level, participated in this study. They were selected from different language institutes in Mashhad in order to remove any probable effect of familiarity with a specific verb use due to instruction of a specific course book. Participants were female students who had studied English at least for a year. They were given a list of 20 sentences; including verbs commonly used erroneously by Persian learners of English, and were asked to translate them into English.

#### *3.2. Materials*

The verbs selected for the study were based on hints in *Common Mistakes in English* and also based on teachers' experience about verbs EFL learners often make mistake about. The translations of verbs, which are used in Persian sentences, were given to students on a piece of paper with a blank in front of them for students to write the English sentences. In order to make subjects produce the desirable sentences the verb to be used was given in parenthesis at the end of each Persian sentence. The sample paper used for data collection is available in Appendix.

#### *3.3. Procedure and analysis*

Pieces of paper were distributed by the researcher. Learners were given twenty minutes to write sentences. The same list of sentences was distributed among intermediate learners to find out if the same kind of errors occurred in learners in higher level of proficiency. The major category of errors



in this study is syntactic one. Sentences containing errors of subcategorization were selected and other kinds of errors were ignored. For example in the sentence *\*We should obey from teacher.* only wrong usage of preposition was considered as error for the study and neglect of article for *teacher* was ignored.

#### 4. Results

Errors were shown within subcategorization frame and the correct selectional restriction of each verb was contrasted with the wrong ones as shown in Table 2:

**Table 2.**

Subcategorizational information

Item	Correct selection	Wrong selection
Agree	verb, [ _ PP(with))]	verb, [to be _ PP(with)]
Answer	verb, [ _ NP]	verb, [ _ PP(to)]
Approach	verb, [ _ NP]	verb, [ _ PP(to)]
Arrive	verb, [ _ PP(at)]	verb, [ _ PP(to)]
Ask	verb, [ _ NP]	verb, [ _ PP(from)/NP PP(from)]
Buy	verb, [ _ NP(someone)NP(something)]	verb, [ _ PP(for someone) NP(something)]
Enjoy	verb, [ _ NP/Reflexive P]	verb, [ _ PP(from)/Reflexive P]
Go	verb, [ _ PP(to)]	verb, [ _ NP]
Hate	verb, [ _ NP]	verb, [ _ PP(from)]
Help	verb, [ _ NP]	verb, [ _ PP(to)]
Let	verb, [ _ NP V]	verb, [ _ NP infinitive]
Many	verb, [ _ PP(to)]	verb, [ _ PP(with)]
Obey	verb, [ _ NP]	verb, [ _ PP(from)]
Reach	verb, [ _ NP]	verb, [ _ PP(to)]
Rob	verb,[ _ NP(the person/place being robbed)]	verb,[ _ PP(from a person or place)]
Say	verb, [ _ PP(to)]	verb, [ _ NP(someone)]
Steal	verb, [ _ NP(the object stolen)]	verb,[ _ PP(from a person or place)]
Tell	verb, [ _ NP]	verb, [ _ PP(to)]
Translate	verb, [ _ PP(into)/NP PP(into)]	verb, [ _ PP(to)/NP PP(to)]
Want	verb, [ _ NP/ infinitive]	verb, [ _ that clause]

Proportion of errors in each level was calculated and demonstrated in Table 3. Evidence produced here suggests that most of the participants were not familiar with verb selectional restriction in the second languages.



**Table 3.**

Proportion of wrong subcategorizational information

Item	Wrong selection	Elementary		Intermediate	
		NO	%	NO	%
Agree	verb, [to be _ PP(with)]	40	100	22	.73
Answer	verb, [_ PP(to)]	34	.80	7	.23
Approach	verb, [_ PP(to)]	38	.95	4	.13
Arrive	verb, [_ PP(to)]	40	100	6	.20
Ask	verb, [_ PP(from)/NP PP(from)]	35	.875	6	.20
Buy	verb, [_ PP(for someone) NP(something)]	31	.775	1	.033
Enjoy	verb, [_ PP(from)/Reflexive P]	29	.725	2	.066
Go	verb, [_ NP]	17	.425	0	.00
Hate	verb, [_ PP(from)]	22	.55	2	.066
Help	verb, [_ PP(to)]	37	.925	10	.33
Let	verb, [_ NP infinitive]	30	.75	3	.10
Many	verb, [_ PP(with)]	40	100	12	.40
Obey	verb, [_ PP(from)]	31	.925	14	.466
Reach	verb, [_ PP(to)]	35	.875	8	.20
Rob	verb, [_ PP(from a person or place)]	40	100	5	.165
Say	verb, [_ NP(someone)]	38	.95	2	.066
Steal	verb, [_ PP(from a person or place)]	40	100	11	.366
Tell	verb, [_ PP(to)]	37	.925	0	.00
Translate	verb, [_ PP(to)/NP PP(to)]	40	100	29	.966
Want	verb, [_ that clause]	26	.65	0	.00

Moreover, comparison of proportion of errors between elementary and intermediate levels shows that percentage of errors decrease as the students improve their language. Therefore, proficiency level is an influential factor.

## 5. Discussion

Learner's performance shown in this study seems to be common in many learners. This could be explained by considering the fact that learners have formed syntactic rules for the words they have acquired either through stages of language learning or transfer from L1 when they find similarities



between L1 and L2. Zobl (1980) stated that in spite of apparently the same structural mechanisms in the origin of both developmental and transfer error, effect of both types of errors on learning would not be the same.

A look at types of errors learners made shows that some errors occurred as a result of L1 interference while some are due to incomplete instruction or imperfect mastery of L2. This is in line with findings of other studies reviewed later in this study (e.g. Abbasi & Karimian, 2011; Behjat & Sadeghi, 2010; Sadeghi, 2009; Namvar et al., 2012). Relatively high percentage of errors made by learners at elementary level makes the necessity of direct teaching of subcategorizational information obvious.

Finding out either L1 or L2 is the source of errors was not the aim of this study. Regardless of the source of errors, it is shown that as proficiency level increased proportion of errors decreases, although not in the same rate. This is also stated by Brown (1994) that a large number of errors happen at the beginning levels of language learning, due to negative transfer and that as language learners improve, intralingual errors overcome interlingual errors.

This indicates that learners need to get familiar with L2 structures through effective instruction in order to avoid errors. Therefore EFL teachers should pay deliberate attention to the explicit teaching of such erroneous terms and provide sufficient opportunities for learners to practice correct usage of verbs both inside and outside the classroom.

It is not only the teachers' responsibility to make learners familiar with accurate verb usage. The problematic vocabularies required to be identified by research findings and clarified in details explicitly in addition to their appearance in natural usage by course book designers and material developers.

## References

- Abbasi, M., & Karimnia, A. (2011). An analysis of grammatical errors among Iranian translation students: Insights from interlanguage theory. *European Journal of Social Sciences*, 25(4), 525-536.
- Behjat, F., & Sadighi, F. (2010). Iranian English major students' L2 grammar development: Linguistic threshold hypothesis. *TESL-HJ*, 13(4). 1-19.
- Brown, H. D. (1987). *Principles of language learning and teaching*. Englewood Cliffs, NJ: Prentice Hall, Inc.



- Brown, H. D. (1994). *Teaching by principles: Interactive language teaching methodology*. New York: Prentice Hall Regents.
- Burt, M. K. (1975). Error analysis in adult EFL classroom. *TESOL Quarterly*, 9(1), 53-63.
- Cook, V. J., & Newson, M. (1996). *Chomsky's universal grammar*. UK: Blackwell.
- Crystal, D. (1992). *An encyclopedic dictionary of language and languages*. Oxford: Blackwell Publishers.
- Dulay, H. C., & Burt, M. K. (1972). Goofing: An indicator of children's second language learning strategies. *Language Learning*, 22, 235-252.
- Fitikides, T. J. (2002). *Common mistakes in English*. England: Pearson Education Limited.
- Khodabandeh, F. (2007). Analysis of students' errors: The case of headlines. *The Asian ASP Journal*, 3(1), 6-21.
- Li, D. C. S., & Chan, A. Y. W. (1999). Helping teachers correct structural and lexical English errors. *Hong Kong Journal of Applied Linguistics*, 4(1), 79-101.
- Namvar, F., Nor, M. F. N., Ibrahim, N., & Mustafa, J. (2012). Analysis of collocations in the Iranian postgraduate students' writings. *The Southeast Asian Journal of English Language Studies*, 18(1), 11-22.
- Nayernia, M. (2011). Writing errors, what they can tell a teacher? *MJAL*, 3(2), 200-218.
- Richards, J. C. (1974). *Error analysis: Perspectives on second language acquisition*. London: Longman Group Ltd.
- Sadeghi, K. (2009). Collocational differences between L1 and L2: Implications for EFL learners and teachers. *TESL Canada Journal*, 26(2), 100-124.
- Sattari, A. (2012). An analysis of grammatical errors in Iranian students' English writings. *Iranian EFL Journal*, 8(2), 143-157.
- Zobl, H. (1980). Developmental and transfer errors: Their common bases and (possibly) differential effects on subsequent learning. *TESOL Quarterly*, 14(4), 469-479.

## Appendix

Translate these sentences using the given verbs in each case.

۱. من با شما موافق هستم. (Agree)
۲. به سوالات زیر پاسخ دهید. (Answer)
۳. گریه آهسته به من نزدیک شد. (Approach)
۴. گریه آهسته به من نزدیک شد. (Arrive)
۵. بچه ها سوالات زیادی ازوالدینشان می پرسند. (Ask)



۶. روز تولدم دوستانم برایم کادو می خرند.(Buy)
۷. من از گوش دادن به موسیقی لذت می برم.(Enjoy)
۸. سال گذشته ساعت هفت صبح به مدرسه می رفتم.(Go)
۹. برادرم از تخم مرغ متنفر است.(Hate)
۱۰. ما در کارهای خانه به هم کمک می کنیم.(Help)
۱۱. اجازه بدھید کمکتان کنم.(Let)
۱۲. برادرم با دختر عمومیش ازدواج کرد.(Marry)
۱۳. ما باید از قانون اطاعت کنیم.(Obey)
۱۴. دستم به لامپ نمیرسد.(Reach)
۱۵. یک گروه دزد از بانک شهر دزدی کردند.(Rob)
۱۶. به مادرت بگو.(Tell)
۱۷. دزد کیفم را زد.(Steal)
۱۸. حقیقت را به مادرم گفتم.(Tell)
۱۹. من همه این جملات را از فارسی به انگلیسی ترجمه کردم.(Translate)
۲۰. معلم از ما می خواهد که به سوالات جواب بدھیم.(Want)

### Bio Data

**Marziye Rabieefar** is an M.A. student at Payam-e-Noor University of Gilan, Iran. She has taught in different English institutes in Mashhad. Her area of interest is error analysis.

**Rajabali Askarzadeh Torghabeh** holds a Ph.D. in English language and literature. He is a faculty member at department of English, Ferdowsi University of Mashhad, Mashhad, Iran. He has published numerous articles in national and international conferences in the areas of English language and literature.