The Strengths and Weaknesses of Iranian IELTS Candidates in Academic Writing Task 2

Reza Panahi

Teaching English as a Foreign Language (TEFEL) Department, Azad University of Shiraz, Shiraz, Iran

Mohammad Mohammaditabar

Teaching English as a Foreign Language (TEFEL) Department, Azad University of Shiraz, Boushehr, Iran

Abstract—This study investigated the strengths and weaknesses of Iranian IELTS candidates in IELTS writing (task 2). This paper used the analytical scoring technique based on the IELTS band descriptor of writing task 2 (public version) 50 writing samples (25 males and 25 females) obtained from four Mock-IELTS tests, taken by IELTS candidates, were used and evaluated by two raters. In order to find the contribution of each component to the total score variance, a multi regression analysis was run. The results showed that Task Response accounted for the largest amount of variance in the total scores (0.429), with Lexical Resource, Grammar Range & Accuracy, and Cohesion & Coherence accounting for 0.345, 0.226, and 0.194, respectively, meaning TR was the strongest point and CC the weakest. Descriptive analysis was run to find the highest scores achieved in each component, which was 5.0 in Task Response, 6.0 in Cohesion & Coherence, 6.0 in Lexical Resource, and 6.0 in Grammar Range & Accuracy. The strengths and weaknesses of candidates in Cohesion & Coherence, Lexical resource, and Grammar Range & Accuracy were investigated using a coding system. First, more than half of the cohesive devices used were accurate. Second, misspelling had the highest frequencies. Last, punctuation was the most counted error. Since CC is the weakest point of Iranian candidates in writing task 2, teachers are advised to address both Cohesion & Coherence and Task Response together since these two criteria are closely related. In terms of Lexical Resource, it is recommended that students be led into programs where they can improve their spelling, as it was the most problematic in the area of LR

Index Terms—task response, cohesion & coherence, lexical resource, grammar range & accuracy

I. INTRODUCTION

According to the official website of IELTS, IELTS is one of the International English Language Testing System, which tests English proficiency across the world. Administering 1.4 million tests all around the world, IELTS has today become the world's most popular English testing system.

IELTS is structurally available in two modules: Academic and General Training. Each is designed to meet particular purposes that candidates have in mind. The IELTS Academic module is designed to evaluate the capability of candidates who needs a communication aspect of language. That is, IELTS, on an international scale, assesses the language proficiency of those who intend to pursue educational purposes in English speaking environments. The IELTS General module intended to fulfill the expectation of candidates who intended to immigrate to English-Speaking countries. These expectations can be undertake work experience or training programs. This module, in other words, is designed for candidates who have vocational intentions in English speaking countries.

IELTS includes tests of all four language skills: Listening, Reading, Writing, and Speaking. The total test length is 2 hours and 44 minutes, made up of listening (30 minutes), Reading (1 hour), Writing (1 hour) and Speaking (11–14 minutes). The Speaking test is administered face-to-face with a trained examiner.

IELTS tests are held in over 500 centers. IELTS removes any diversity and discrimination to everyone who sits the test, regardless of their nationality. IELTS has a trustable quality and high security due to three reputable organizations which are *British Council, IDP: IELTS Australia and the University of Cambridge ESOL Examinations (Cambridge ESOL).* Due to the high quality and controlled security procedure, lots of governments and universities rely on it.

Purpose of the Study

The reason IELTS writing module is chosen as the focus of this research is many-fold. In the first place, the writing subtest of IELTS seems to be Achilles' heels to every candidate due to many reasons, and many lose scores in writing and consequently gain a low overall band score. It seems necessary that candidates' performance be scrutinized based on IELTS marking system and under simulated IELTS examination conditions in order to have a better understanding of the blocking problems causing the yield of low scores. These problems appear to be in direct relation with the conditions under which the test is taken and the writing grading system of IELTS. The public version of this marking scheme is under study in this paper. By applying the standards, pre-designed by ILETS officials, in this band descriptor, found in the appendix, the major issues of Iranian IELTS candidates in writing task 2 are expected to be elicited.

Research questions

Based on the objectives, this study seeks to answer the subsequent questions:

1- What is the relative weight for the variance of each component that contributes to the overall writing proficiency scores of Iranian students?

2- What is the percentage of successful addressing of the task based on IELTS band descriptor?

3- What is the percentage of effective use of cohesion based on IELTS band descriptor?

4- What is the percentage of effective use of lexical items based on IELTS band descriptor?

5- What is the percentage of the accurate and inaccurate use of the four forms of sentences: simple, compound, complex, and compound complex?

II. METHODOLOGY

Participants

The participants of this study were 50 candidates (25 male and 25 female) of an IELTS House in the south of Iran. They were selected based on convenience or opportunity sampling procedures. This is because a similar sampling close to that of a real IELTS test is the underlying intention of the researcher.

These candidates were all Persian speakers, prepared to take part in a real IELTS test in the future.

Instrumentation

As cited in ielts.org (IELTS | Researchers - Band descriptors, reporting and interpretation, 2012) examiners award a band score for each of four criterion areas: Task Achievement (for Task 1), Task Response (for Task 2), Coherence & Cohesion, Lexical Resource & Grammatical Range and Accuracy. The four criteria are equally weighted on the scale of 1 to 9. The modified analytical style of the International English Language Testing System (IELTS) writing scales were used (Shaw, 2002). Thus, the accessibility of more comprehensive descriptions of written language ability at each band level appears highly advantageous. Some key features such as the qualities of learners' performance, the accuracy of their performance during the task, and distinguishing all the band levels considered as important elements in assessing process.

Understanding the essential qualities at any different level, will help one to comprehend the L2 writing task better (Weigle, 2002; Hawkey and Barker, 2004). Also knowing the linguistic descriptions were turn out to be one of the effective feature of writing task which discriminate one level of performance from the other one. Such an account would also allow test-makers to make descriptors more detailed. This would be well received by IELTS raters" (Shaw, 2004).

Data Collection Procedure

The focus of the present study was the IELTS Academic Writing subtest. In this module, in a real IETS Test, candidates were given two writing tasks, the second of which was under study in this paper. In Writing Task 2, candidates were given a topic such as the ones below to write about:

> WRITING TASK 2
> You should spend about 40 minutes on this task.
> Write about the following topic:
> **Children who are brought up in families that do not have large amounts of money are better prepared to deal with the problems of adult life than children brought up by wealthy parents.**
> To what extent do you agree or disagree with this opinion?
> Give reasons for your answer and include any relevant examples from your own knowledge or experience.
> Write at least 250 words.

> WRITING TASK 2
> You should spend about 40 minutes on this task.
> Write about the following topic:
> **Some people believe that international tourism has brought enormous benefit to many places. On the other hand, there is concern about its impact on local inhabitants and the environment. Discuss both views and give your opinion.**
> Write at least 250 words.

40 minutes were given to the candidates. They were asked to write at least 250 words. Candidates should write well-organized, relevant, and to the point. They can support their ideas by giving examples or evidence.

In this task, candidates confront two types of tasks: a discussion and an argument task. Candidates are required to write an argument essay on the latter using their own opinions and use their own experiences to support the main topic. In the former type, a discussion essay needs to be written where each of the two given views are discussed without the interference of the candidate's personal views. In the conclusion, the writer will comment which view is acceptable to him or her. The writer can reasonably accept either ones, both, or none of the views provided the choice be supported.

Candidates receive scores on a Band Scale from 1 to 9. A profile score is reported for each skill. The four individual scores belonging to each language skill are averaged and rounded to produce an Overall Band Score.

In IELTS, each task is assessed independently. Detailed performance descriptors have been developed which describe written performance at the nine IELTS bands. Public versions of these descriptors are available on the IELTS website (www.ielts.org). The descriptors are based on the following criteria (for task 2 only):

• Task Response (TR)
• Coherence & Cohesion (CC)
• Lexical Resource (LR)
• Grammatical Range & Accuracy (GRA)

Procedure

In this study, the candidates were IELTS intending trainees who attended Mock-IELTS. Mock-IELTS (MI) is normally held twice a semester in the alleged institute (every forty-five days), and all IELTS students will have to sit for this test to see the result of five weeks of preparation for the real test. None of the candidates knew their work was going to be analyzed as this information could jeopardize the integrity of this study. Only when the MI was over, everyone was informed of the process to which every individual consented.

On all four MIs, similar procedures as in a real IELTS test were applied. In the same manner, all ID cards were checked. Cell-phones and extra belongings were collected. In the exam area, Farsi was not allowed. Introductory speech on the dos and don'ts of the test was given by the researcher. The test began at a certain time starting with listening, and then reading. Finally, with writing, the same amounts of introductory guideline as in IELTS were given. Note that the researcher had sat for the test of IELTS three times, achieving band 8 in all three tries. First the answer sheets, and then the questions' booklets were handed out. The booklets had the same cover as in a real IELTS test. The answer sheets were also the original answer sheet of IELTS writing adopted from IELTS Official Materials (2009). Having finished the tests, all the papers were collected and all the parts of the test except for writing task 2 were excluded.

Data analysis

The data collected from the scorings and analyses above were transferred to IBM SPSS statistics 19 for statistical computations and analysis. The significance level was set at 0.05. The following analyses, afterwards, respectively were done.

1) Because the researcher scored the papers twice with ten days interval, intra-rater reliability needed to be estimated by calculating a correlation coefficient between the two sets of scores (Brown, 2005).

2) Inter-rater reliability between the final set of scores of the researcher calculated above and the set of scores presented by the second rater was calculated in the same manner followed in step 1.

3) A linear regression analysis via SPSS was calculated to estimate the relative weight for the variance of each component (criterion) that contributes to the overall writing proficiency scores of Iranian students.

4) Independent samples *t*-test in SPSS was applied to see if there was a relationship between age, education, and language competence of candidates and the overall score they have achieved.

5) The coded errors were compared in terms of their frequencies to show the sources of strenghs and weaknesses that Iranian IELTS candidates have.

III. RESULTS AND DISCUSSION

In this section, the researcher covered all the claculations and estimations leading to the relative significance for each of the varience of each criterion (out of the four criteria of IELTS in assessing written performances) contributing to the overal writing profeciency scores of Iranian IELTS candidates are presented. Then the evaluating compositions (task 2) are statistically demonstrated and the effect of gender, educational background, ange and language competence of the participants on the outcome of their written performances were measured.

The relative weight for the variance of each criterion contributing to the overall writing proficiency scores

In the first place, the descriptive characteristics were calculated and the results are summarized in Graph 1 illustrates the three ratings of the 50 essays in the present study. By looking at the bar charts, it is easily seen that both examiners had similar ratings except for TR where the researcher's colleague seems to have a dissimilar view towards the organization of an essay.
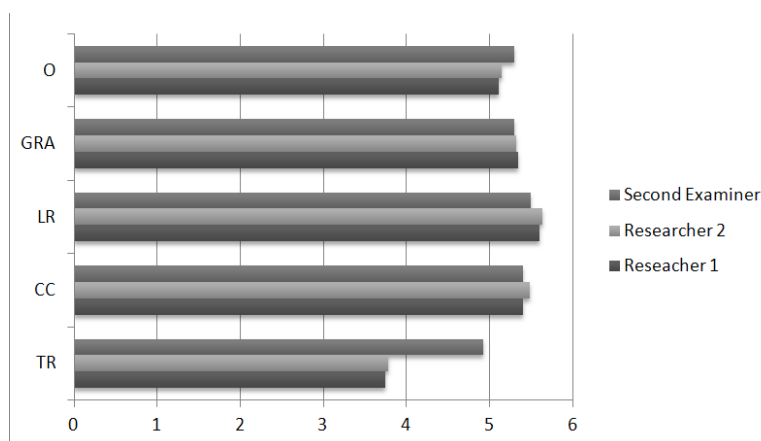
Graph 1: Comparisons of the three scorings

Intra-rater reliability

The results of intra-rater reliability are summarized in Table 1.

TABLE 1:
INTRA-RATER RELIABILITY

	Criteria	Pearson Correlation	Spearman-Brown Prophecy
1	TR	0.900	0.95
2	CC	0.708	0.83
3	LR	0.888	0.94
4	GRA	0.683	0.81
5	Overall	0.862	0.93

According to Cohen (1988) correlation scores between the scales of 0.1 up to 0.29 are considered as small correlations, between the scales of 0.30 up to 0.49 are medium correlations, and correlation scores between the scales of 0.50 to 1 are large correlations that show high inter-relations. This table indicates that the intra-rater reliability between the two scorings of the researcher is high because it falls in Cohen's third category. The results of Inter-rater reliability also showed with the highest amount for LR (0.80) and the lowest for GRA (0.64).

In order to determine the amount of variance that is contributed by each composition component to the total scores a multiple linear regression analysis was run. The purpose was to determine how productive or significant each component was and how much variance was accounted for by each component. In this way, the weakest and strongest areas of candidates' performance were identified.

Plotting the data of the dependent variable against each independent variable was run to find the linear relationship between the dependable and independable variables. (Norusis, 1988). The results show that the relationships between the dependent variable (total scores) and the independent variables (components) were all reasonably linear. Having run the regression analysis by SPSS, the following results were obtained.

TABLE2:
MODEL SUMMARY

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.975[a]	.951	.946	.1491

In Table 2, it is clear that the coefficient of multiple determination is 0.951; therefore, about 95% of the variation in the final scores is explained by the four criterion. The regression equation appears to be very useful for making predictions since the value of R^2 is close to 0.1.

TABLE 3:
ANOVA

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	19.305	4	4.826	217.205	.000[a]
	Residual	1.000	45	.022		
	Total	20.305	49			

Table3, reveals that at $\alpha = 0.05$ level of significance, it is clear that at least one of the predictors is useful for predicting the overall scores; therefore the model is useful.

TABLE 4:
COEFFICIENTS

Model		Standardized Coefficients	t	Sig.	Correlations	
		Beta			Zero-order	Partial
1	(Constant)		0.39	.699		
	TR	0.429	10.70	.00	0.78	0.84
	CC	0.194	3.84	.00	0.80	0.49
	LR	0.345	6.79	.00	0.84	0.71
	GRA	0.226	4.68	.00	0.74	0.57

Table 4 shows that At $\alpha = 0.05$ level of significance, there exists enough evidence to conclude that the slopes of LR, CC, LR, and GRA variables are not zero and, hence, the scorings are useful predictors of overall scores.

Based on what is shown is Table 4, the first essential question of this study is answered. As it can be seen TR (β=0.429) is the strongest predictor of the total score, meaning Iranian IELTS candidates' strongest point in their written performance in the test of IELTS is Task Response, which is the ability to develop a position to a given prompt in the form of a question or statement. On the other hand, Cohesion & Coherence criterion (β=0.194), the overall clarity and fluency of the message, is the weakest area in Iranians' writing productions. That is Iranians do not seem to be able to present the message of their statements clearly and fluently. This part of the analysis is completely in line with the researcher's expectations.

Finally, the second and the third strong points in a descending order are LR (β=0.345) and GRA (β=0.226).

To recap, TR, LR, GRA, and CC, respectively and in a descending order, cover the range of writing skills demonstrated in task 2 of IELTS by IELTS intending Iranian candidates.
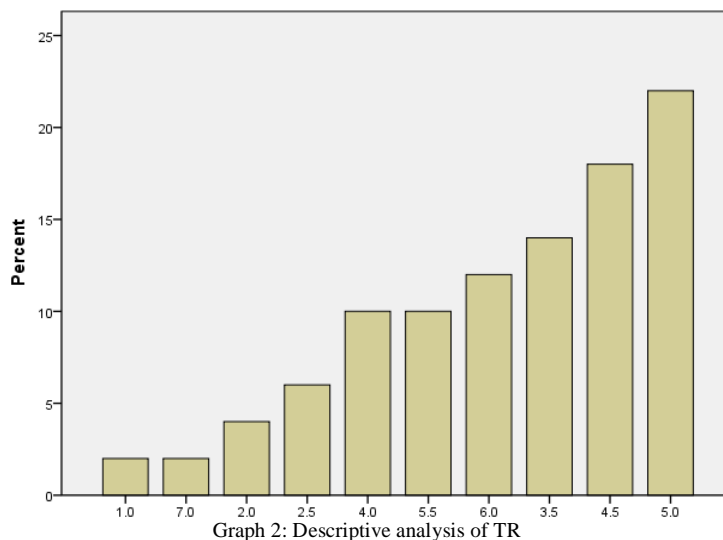
TABLE 5:
COEFFICIENTS

Model		Correlations	Co linearity Statistics	
		Part	Tolerance	VIF
1	(Constant)			
	TR	0.354	0.681	1.469
	CC	0.127	0.431	2.319
	LR	0.225	0.425	2.355
	GRA	0.155	0.472	2.117

Table 5 demonstrates that since none of the predictor variables has a variance inflation factor (VIF) greater than ten, there are no obvious multi-collinearity problems; in other words, there is no variable in the model that is measuring the same relationship as is measured by another variable or group of variables.

To answer the second question that this study is concerned with, the data in the following table (6) and graph (2) are presented.

TABLE 6:
DESCRIPTIVE ANALYSIS OF TR

N	Valid	50
	Missing	0
Mean		4.46
Std. Error of Mean		.171
Mode		5.0
Std. Deviation		1.21
Variance		1.47
Range		6.0
Minimum		1.0
Maximum		7.0
Sum		223.

Graph 2: Descriptive analysis of TR

As it can be seen in Table18, the minimum and maximum scores are achieved by only 3% of the participants as shown in Graph 2. This means that only one participant's performance was completely unrelated and one participant's response addressed all parts of the task, and presented a clear position throughout the essay.

Band 6 is the range, which comprises only about 12% of participants who addressed all parts of the task although some parts may be more fully covered than others. Band 5.5 is achieved by 10 %. However, over 22% of participants achieved band 5, the most frequently achieved band. These participants generally addressed the task only partially and expressed a position but the development was not always clear. Bands 4 and 4.5 are achieved by approximately 10% and 14%, respectively. These candidates responded to the task only in a minimal way, and their position was not clear. Bands 2 and 2.5 belong to an average of 5% of participants who barely responded to the task and expressed no position

Effective use of CC

In terms of CC, the third question, the following tables and graphs were developed.

TABLE7:
DESCRIPTIVE ANALYSIS OF CC

N	Valid	50
	Missing	0
Mean		5.510
Std. Error of Mean		.0984
Mode		6.0
Std. Deviation		.6961
Variance		.485
Range		3.0
Minimum		3.5
Maximum		6.5

Graph 3: Descriptive analysis of CC
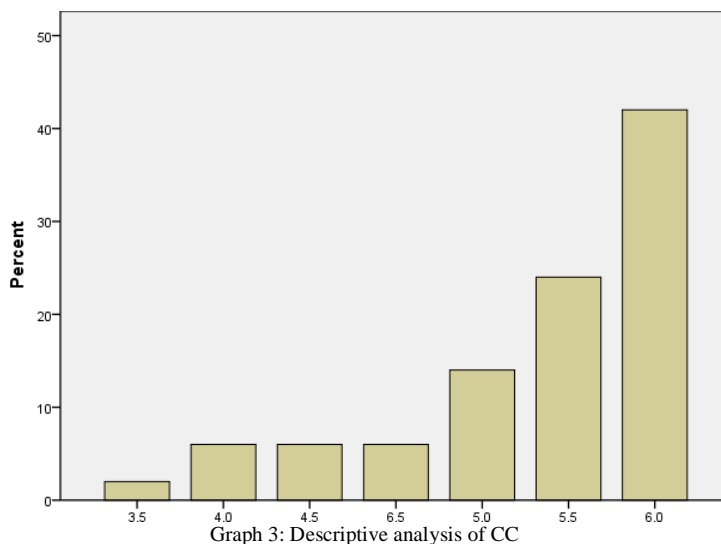
The minimum and maximum band-scores (Table7) in this criterion are 3.5 and 6.5, given to only an average of 4% of candidates as Graph 3 reveals. The highest score achieved is 6, which means around 40% of candidates arranged information and ideas coherently, and there is clear overall progression in their performance. Approximately, 15% to 25% of participants achieved bands 5 and 5.5. They presented information with some organization, but there may be a lack of overall progression in their compositions. Band-scores 4 and 4.5 are obtained by only about 5% of examinees. These candidates presented information and ideas not coherently and there is no clear progression in their work.

TABLE 8:
DESCRIPTIVE ANALYSIS OF ERRORS IN CC

|		CC10 NEG	CC10 POSI	CC10 TOTAL	CC20 NEG	CC20 POSI	CC20 TOTAL
N	Valid	50	50	50	50	50	50
	Missing	0	0	0	0	0	0
Mean		4.44	11.66	16.10	4.38	9.32	13.70
Std. Error of Mean		.610	.773	.871	.379	.482	.604
Mode		2	10	12[a]	4	11	15[a]
Std. Deviation		4.315	5.465	6.162	2.679	3.407	4.273
Variance		18.619	29.862	37.969	7.179	11.610	18.255
Range		23	23	23	10	18	21
Minimum		0	1	6	0	1	4
Maximum		23	24	29	10	19	25
Sum		222	583	805	219	466	685

Table 8 shows out of 805 referential cohesive devices used in all papers, 222 of them were incorrect and 583 of them were correct. Moreover, in each composition an average of 4.44 incorrect and 11.66 correct referential devices (out of an average 16.10) was found.

All in all, 685 connective cohesive devices were used in all 50 papers. 466 correct instances of the correct use of these words were recorded, which left only 219 incorrect use all together. Out of 13.70 connective cohesive devices in each paper, an average of 4.38 incorrect and 9.32 correct items were spotted. In conclusion, candidates were more accurate in terms of CC than inaccurate. The analysis in this section is visually summarized in Graph 7.

Effective use of LR

TABLE 9:
DESCRIPTIVE ANALYSIS OF LR

N	Valid	50
	Missing	0
Mean		5.590
Std. Error of Mean		.1027
Mode		6.0
Std. Deviation		.7263
Variance		.527
Range		3.0
Minimum		4.0
Maximum		7.0
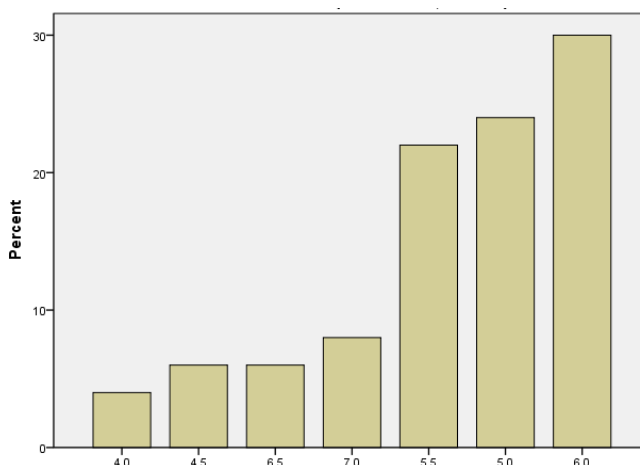
Graph 4: Descriptive analysis of LR

Looking at the LR in Table 8, it is easily visible that minimum and maximum band-scores are 4.0 and 7.0, achieved by an average of only 5% of candidates, based on Graph 4.8. Band scores 6.0 (the highest score achieved) and 6.5 are obtained by 5% to 30% of candidates. These participants used an adequate range of vocabulary for the task and attempted less common lexical items with some inaccuracies. Bands 5.0 and 5.5 are given to an average of 21% because they used a limited range of vocabulary, which was minimally adequate for the task.

TABLE 9.1.
DESCRIPTIVE SCRUTINY OF DIFFERENT AREAS IN LR

		LR NEG01	LR POSI01	LR02	LR03	LR NEG05	LR POSI05
N	Valid	50	50	50	50	50	50
	Missing	0	0	0	0	0	0
Mean		2.02	0.00	7.42	1.38	0.68	0.38
Std. Error of Mean		0.238	0.000	0.818	0.249	0.129	.099
Median		2.00	0.00	6.00	1.00	0.00	0.00
Mode		0[a]	0	3	0	0	0
Std. Deviation		1.684	0.000	5.782	1.760	0.913	0.697
Variance		2.836	0.000	33.43	3.098	0.834	0.485
Range		6	0	23	7	3	2
Minimum		0	0	0	0	0	0
Maximum		6	0	23	7	3	2
Sum		101	0	371	69	34	19

TABLE 9.2.
DESCRIPTIVE SCRUTINY OF DIFFERENT AREAS IN LR

		LR06	LR07	LR08
N	Valid	50	50	50
	Missing	0	0	0
Mean		5.40	2.44	1.54
Std. Error of Mean		0.590	0.323	0.268
Median		4.00	2.00	1.00
Mode		1	0	0
Std. Deviation		4.170	2.287	1.897
Variance		17.38	5.231	3.600
Range		20	8	7
Minimum		0	0	0
Maximum		20	8	7
sum		270	122	77

According to Tables 9.1and 9.2, the most serious error committed is spelling; each composition contains an average of 7.42 misspellings, and all together, 371 misspellings were counted. The next large mean belongs to uncommon lexical items with an average of 5.40 in each writing. A sum of 270 new vocabularies were used by candidates. In addition, an average of 2.44 wrong words was detected in performances, and 2.02 incorrect word formations in average exist in essays. The statistics belonging to the rest of inaccuracies and accuracies were not as significant. Range and accuracy of grammar

TABLE 10.
DESCRIPTIVE ANALYSIS OF GRA

N	Valid	50
	Missing	0
Mean		5.380
Std. Error of Mean		.0722
Mode		5.0
Std. Deviation		.5107
Variance		.261
Range		2.0
Minimum		4.0
Maximum		6.0
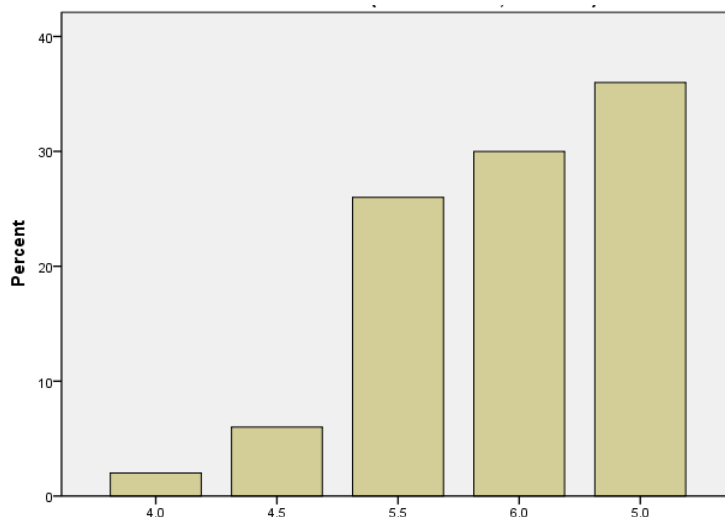
Graph 4. Descriptive analysis of GRA

As can be seen in Table 10, the minimum and maximum band-scores of GRA are 4.0 and 6.0. Less than 2% and more than 5%, according to Graph 4, achieved 4.0 and 4.5, meaning that they used only a limited range of structures with only rare use of subordinate clauses. Band 6.0 belongs to nearly 30% of participants who used a mix of simple and complex sentence forms. Band 5 and 5.5 were given to around 60% of candidates because they used only a limited range of structures and attempted complex sentences but those tended to be less accurate than simple ones. This includes well above half of the candidates in this study.

What follows next is the detailed scrutiny of the performances of GRA summarized in the following tables.

TABLE 11.1.
DESCRIPTIVE ANALYSIS OF ACCURACIES AND INACCURACIES IN GRA

		GRA 001NEG	GRA 001POSI	GRA001 T	GRA 002NEG
N	Valid	50	50	50	50
	Missing	0	0	0	0
Mean		1.68	3.16	4.84	0.78
Std. Error of Mean		0.264	0.320	0.362	0.125
Median		1.00	3.00	5.50	1.00
Mode		0	2	6ᵃ	0
Std. Deviation		1.867	2.262	2.558	0.887
Variance		3.487	5.117	6.545	0.787
Range		7	8	11	3
Minimum		0	0	0	0
Maximum		7	8	11	3
Sum		84	158	242	39

TABLE11.2.
DESCRIPTIVE ANALYSIS OF ACCURACIES AND INACCURACIES IN GRA

		GRA 002POSI	GRA002T	GRA 003NEG	GRA 003POSI
N	Valid	50	50	50	50
	Missing	0	0	0	0
Mean		0.62	1.40	2.90	2.60
Std. Error of Mean		0.137	0.176	0.265	0.291
Median		0.00	1.00	3.00	2.00
Mode		0	1	2ª	2
Std. Deviation		0.967	1.245	1.876	2.060
Variance		0.934	1.551	3.520	4.245
Range		4	5	9	10
Minimum		0	0	0	0
Maximum		4	5	9	10
Sum		31	70	145	130

TABLE 11.3.
DESCRIPTIVE ANALYSIS OF ACCURACIES AND INACCURACIES IN GRA

		GRA 003T	GRA 004NEG	GRA 004POSI	GRA 004T	GRA 005
N	Valid	50	50	50	50	50
	Missing	0	0	0	0	0
Mean		5.48	0.66	0.50	1.16	9.52
Std. Error of Mean		0.313	0.127	0.115	0.157	0.817
Median		5.50	.00	.00	1.00	8.50
Mode		6	0	0	0	5
Std. Deviation		2.215	0.895	0.814	1.113	5.779
Variance		4.908	0.800	0.663	1.239	33.38
Range		10	4	3	4	22
Minimum		1	0	0	0	1
Maximum		11	4	3	4	23
Sum		274	33	25	58	476

TABLE 11.4.
DESCRIPTIVE ANALYSIS OF ACCURACIES AND INACCURACIES

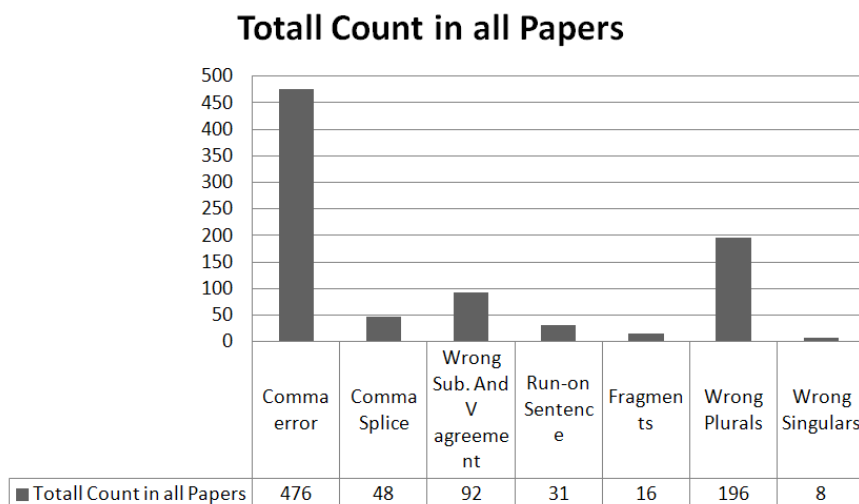|		GRA 006	GRA 007	GRA 008	GRA 009	GRAPL	GRASIG
N	Valid	50	50	50	50	50	50
	Missing	0	0	0	0	0	0
Mean		0.96	1.84	0.62	0.32	3.92	0.16
Std. Error of Mean		0.298	0.251	0.166	0.088	0.567	0.072
Median		0.00	1.00	0.00	0.00	2.00	0.00
Mode		0	1	0	0	0ª	0
Std. Deviation		2.109	1.777	1.176	0.621	4.009	0.510
Variance		4.447	3.158	1.383	0.385	16.07	0.260
Range		12	7	6	2	16	3
Minimum		0	0	0	0	0	0
Maximum		12	7	6	2	16	3
Sum		48	92	31	16	196	8

Based on Table11.1, 242 simple sentences were used in all 50 essays. Out of these, 158 sentences (an average of 3.16 in each paper) were accurate and only 84 ones (an average of 1.68 in each) were inaccurate. Compound sentences, according to Tables 11.1 and11.2 were 70, out of which 39 were inaccurately used and 31 accurately. Candidates used complex sentences the highest: 274 instances (Tables11.2 and11.3). Out of these, an average of 2.90 were used inaccurately in each paper (145 inaccurate all together) and an average of 2.60 (130 in total) accurately. Compound-complex sentences were used the lowest: only 58 times, most of which (33) were used incorrectly.

Comma splice, run-on sentences, and fragments are errors commonly found in candidates' writings. In the first one, two sentences are joined with just a comma. Among these compositions 48 instances were found according to Table 11.4. The second error (connecting two sentences using no punctuation) had a lower figure: 33 times. Fragments or sentence-like phrases had the lowest figure, which were only 16 cases.

In terms of punctuation, as shown in Table 11.3., 476 instances of inaccurate use of comma were detected, which accounts for an average of 9.52 in each paper. This area seems to be highly problematic to participants.

Lastly, as shown in 11.4, the error related to subject-verb agreement was seen 92 times in all papers. This equals an average of 1.84 in each paper. Singular words that had to be plural, as seen in 11.4, had a high figure: 196 cases, an equal of 3.92 in each written performance

In Graph 5, the summary of all counted errors in GRA is presented for a better understanding of the statistics in this section.

Totall Count in all Papers

	Comma error	Comma Splice	Wrong Sub. And V agreement	Run-on Sentence	Fragments	Wrong Plurals	Wrong Singulars
■ Total Count in all Papers	476	48	92	31	16	196	8

Graph 5. Summary of counted errors in GRA

IV. CONCLUSION AND IMPLICATIONS OF THE STUDY

As it was clear the weight of each IELTS writing component (TR, CC, LR, and GRA) on the final score (Overall) was measured, and it was concluded that what Iranian IELTS candidates show more dexterity towards, is the use of Task Response. TR is the ability to formulate and develop a position in relation to a given prompt in the form of a question or statement. The weakest point for this group was Cohesion & Coherence. As it was seen, Iranian IELTS candidates may not be able to create an essay in which the ideas are connected coherently and cohesively. They seem to have difficulty writing a composition, which shows a clear progression throughout. LR and GRA are the next criteria affecting the overall score, respectively.

Then the successful addressing of the task (TR) was analyzed. The minimum score was one and maximum 7. The most achieved band was band-score 5.

the effective use of Cohesion & Coherence was analyzed. The minimum score was 3.5 and the maximum 6.5. The highest band obtained was 6.

the skillfulness of candidates in using lexical items was scrutinized. The minimum score in LR was 4, and max 7. The highest band achieved was 6.

GRA was put to scrutiny. The lowest score was 4.0 and the highest 6.0. The highest score in GRA was 5.

Based on the result of the present study, unlike what was expected by the researcher, Iranian IELTS candidates appear to be more skillful at putting together a position and addressing all parts of the given task (TR).

Following this strong point, LR and GRA are deemed the decisive factors in determining one's total score in IELTS writing (task 2), respectively. Finally, creating a cohesive and coherent paper in which ideas are well connected is not easy for Iranians (CC), and appears to be, based on the calculations in this study, the weakest area in writing subtest of IETS for Iranian IELTS candidates.

Implications of the study

The implications of this study are many-fold. First of all, in preparing students for the test of IELTS, particularly for writing task 2, the common practice for teachers is that they have students write an essay, and then they red-pen the whole essay, commenting on grammatical errors. However, what is recommended based on the findings in this study is that grammar range & accuracy is related to candidates' prior implicit and explicit knowledge of grammar that cannot be effectively improved in such short courses like IELTS preparation classes. Instead, teachers are advised to raise students' consciousness towards the four main sentence structures: simple, compound, complex, and compound-complex. Teachers should also raise candidates' awareness towards the three main errors in writing, namely run-on, comma splice, and fragments. Note that raising candidates' awareness in short courses like IELTS, based on this study, might be more effective rather than taking measures to teach students such things.

TR and CC should be addressed together since IELTS intending students deem to feel difficulty with CC as it was the weakest point in this paper. Having learnt to take a stand and formulate a clear position that includes all aspects and parts of the task (TR), students should learn to organize their supporting ideas logically so that their essays display a natural progression throughout the composition (CC). Focusing on CC will help students to learn the four major cohesive devices: Conjunctive adverbs (e.g. however), coordinating conjunctions (e.g. and), relative pronouns (e.g. who) and subordinate conjunctions (e.g. as soon as). It also inculcates with them the importance of referencing (e.g. use of pronouns) more effectively. Teachers should also clarify that over or under-use of these devices may result in loss of scores. This can be done by showing candidates concrete examples.

LR in this study was plagued with misspellings. Teachers should guide students to take recommended exercises to improve their spelling. Candidates should also be discouraged to use any "big" words that they randomly learn, as they may appear inappropriate or wrong.

The implications of this study concerns material-developers as well. Books compiled or written about writing task 2 should familiarize students with all four criteria that compositions are measured against. Clear examples should be provided to illuminate ways that help candidates raise their score. Every sample printed in such books should come with an analysis of the content of the essay in relation to the four criteria. In this way, candidates' consciousness raises effectively, and at the time of writing candidates will probably act wisely in writing the response to the task.

REFERENCES

[1]   Brwon, J. D. (2005). Testing in Language Programs. New York: MacGraw-Hill ESL/ELT.
[2]   British Council, Cambridge ESOL, and IDP: IELTS Australia. (2009). Official IELTS Practice Materials. Cambridge: University of Cambridge. ESOL Examinations, IELTS Australia, British Council - University of Cambridge, ESOL Examinations.
[3]   Cohen, J. (1988). Statistical Power Analysis for the Bahavioral Science (2nd ed.). Hillsdale, NJ: Lewrence Earlbaum Associates.
[4]   IELTS | Researchers - Band descriptors, reporting and interpretation, 2012. (2012, 6 22). Retrieved 6 22, 2012, from www.ielts.org: http://www.ielts.org/researchers/score_processing_and_reporting.aspx.
[5]   Harrison, M., & Whitehead, R. (2006). IELTS Exam Essentials: IELTS Practice Tests. London: Christopher Wenger.
[6]   Hawkey, R., & Barker, ,. F. (2004). Developing a common scale for the assessment of writing. *Assessing Writing 9* (2), pp 122-59.
[7]   Norusis, M. (1988). PSS/PC+Studentware. Chicago: IL:SPSS Inc.
[8]   Official IELTS Materials. (2009). The British Council, Cambridge ESOL, and IDP: IELTS Australia. University of Cambridge | 2009 | ISBN-10: 1906438463 | ISBN-13: 978 1906438463.
[9]   Shaw, S. (2004). 'IELTS writing: revising assessment criteria and scales (Phase 2). *Research N otes*, 15, 9-11.
[10] Weigle, S. (2002). Assessing writing . Cambridge: Cambridge University Press. Retrieved 3 23, 2012, from www.ielts.org.

Reza Panahi got his M.A. in TEFEL (Teaching English as a Foreign Language (in Shiraz Azad University, Fars, Iran in October2012. He was born in Shiraz. He teaches English more than 10 years and he has done his military service. Mr. Panahi has not been the member of any professional societies. He has also published another article under the name of the Effects of Task Type Planning on Students' Essay Writing: A Study of Iranian EFL Learners.

Mohammad Mohammaditabar holds an MA in TEFL. He is currently the President of Safirelian Language and IELTS House, Bushehr, Iran. Mohammaditabar has published a book that serves as a handbook for teacher training courses in Iran. His main interests are international exams, and theories of language teaching and learning.