Task Types in EFL Context: Accuracy, Fluency, and Complexity in Assessing Writing Performance

Natasha Pourdana and Sayyed Mohammad Karimi Behbahani

Abstract—**Inspired by the recent surge of interests in EFL writing assessment, this study intended to examine the impact of three types of language assessment tasks - Topic Writing, Picture Description, and Text Reconstruction - on aspects of accuracy, fluency and complexity in Iranian EFL learners' writing performance. Findings from Analysis of Variances (ANOVA) demonstrated a high degree of accuracy and complexity in EFL learners' performance on Topic Writing task, comparing to aspect of fluency which was demonstrated to be the highest in EFL learners' performance on Picture Description task. Suggestions for simultaneously improving various aspects of EFL learners' writing performance were also provided.**

Index Terms—**Accuracy, assessment tasks, complexity, fluency, writing,**

I. INTRODUCTION

Over the last twenty years, communicative tasks have evolved as an important component within curriculum planning, implementation, and evaluation in language teaching and learning. In task-based language teaching, syllabus content and instructional processes are selected with reference to the communicative tasks in which language learners need to engage in and outside the classroom; also with reference to the approaches and empirical insights into those social and psycholinguistic processes which facilitate language acquisition.

As Skehan [1] states, a communicative task is an activity in which, (a) meaning is primary, (b) there is some sort of relationship to real world tasks (c) task completion has some sort of priority, and (d) assessment of task performance is determined in terms of task outcomes. Similarly, Nunan [2] believes that a task is a classroom work that engages learners in completing, manipulating, producing, or interacting in the target language while the attention is focused on meaning rather than on form. Also, Crabbe [3] asserts that task can provide a framework for communicative performance. Therefore, behind every task there is a set of learning opportunities and potential activities for learning.

EFL teachers can make use of different tasks as teaching different materials in their classrooms. Using tasks would be

Natasha Pourdana is a PHD holder in applied linguistics and a full-time assistant professor in Department of English Translation, Islamic Azad University, Karaj Branch, Iran (e-mail: Natasha.qale@gmail.com).

Sayyed Mohammad Karimi Behbahani is a PHD holder in English translation studies and a full-time assistant professor in Department of English Translation, Islamic Azad University, Karaj Branch, Iran (e-mail: m.behbahani@kiau.ac.ir).

beneficial in teaching and assessing writing performance, too. Creating a new and different situation for language learners, communicative tasks will help them use their abilities to solve language problems in doing tasks; hence language learning experience would be easier and more interesting [2]. Such a communication in a second or foreign language is a highly complex but fascinating activity.

As Richards and Renandya [4] believe, there is no doubt that writing is the most difficult skill for second language learners to master. The difficulty lies not only in generating and organizing ideas but also in translating those ideas into meaningful texts. Tasks are, therefore, fundamental in learning to write and represent a central aspect of the teacher's planning and delivery of a writing course. The tasks teachers assign will help students to learn from their experience, to develop an understanding of the text and to control their writing skill.

Writing needs practicing and internalizing a set of structures that can promote a balanced development of learners' fluency, accuracy, and complexity in the target language. As Skehan [1] emphasizes, "the more the task is planned, the less computational work needs to be done during the task performance. Things being equal, the result is more, when attention is given as a general tool to achieve a variety of goals such as greater fluency, accuracy, and complexity" (p. 73). He also distinguishes three aspects of linguistic performance: (a) Fluency, which is concerned with the learners' capacity to produce language in real time without any pauses or hesitations. Fluency is measured in different ways, such as speech rate, length of the run, pause length, false starts repetitions, and reformulating; (b) Accuracy, which is the extent to which the language produced conforms to the target language norms. There are different researches to accuracy either as an error free piece of language or accurate use of specific forms (Skehan and Foster, 1997); and, (c) Complexity which is the elaboration of the language produced by EFL learners. Researchers usually measure complexity by relating how much subordination are used per T-units or C- units.

With gradual acceptance of errors as productive and developmental rather than substandard and deviant, grammatical accuracy became secondary to communication. English L2 composition text books reflected the theoretical shift by focusing on the teaching of organization patterns common in English academic prose, topic, thesis sentences, paragraphs, and essay modes with their focus primarily on product. This current traditional approach is still widely used in many writing classes.

At the beginning of the 21st century, writing classrooms have achieved a more balanced perspective of composition

theory. Consequently, new pedagogy has begun to develop traditional teachers-centered approaches into more learner-centered courses so academic writing was viewed as communicative social act.

Based on a widely accepted categorization, writing tasks in second/foreign language classes are either *real-world* tasks, which are directly based on the learners' communicative goals, or *pedagogic* tasks which are designed to develop students' genre knowledge and composing skills.

Many pedagogic tasks aim to promote discrete skills, such as improving punctuation, developing pre-writing abilities, or increasing an understanding of rhetorical forms. These tasks are selected on the basis of meta-cognitive criteria, or what students need to know in order to build the competence required to accomplish real-world objectives on later stages [3].

Pedagogical tasks provide a vehicle for the presentation of appropriate target language samples to learners and for the delivery of comprehension and production opportunities of negotiable difficulty. Murphy [5] emphasizes the fact that communicative tasks may be chosen and implemented so that particular pedagogical outcomes are achieved. Such tasks should carefully be designed to lead students to the intended objectives. He also distinguishes among the factors that affect learning outcome, contribution of individual learner, the task performance, and the situation in which the task is performed.

In topic writing tasks, for example, students are required to write free compositions on carefully chosen realistic topics. Composition can be a useful writing assessment task, too. "Topic writing tasks provide students with an opportunity to demonstrate their ability to organize language materials, using their own words and ideas to communicate. In topic writing students should be presented with a carefully defined problem which motivates them to write. They should also have an audience in mind when they write" [6].

Similarly, when students try to describe something they are usually trying to give the reader an exact and detailed impression of something in their minds. In description writing, they are usually concerned with allocating such features as place, position, direction, measurement, weight, size, volume, distance, and shapes and patterns" [7]. Since, most descriptions are mostly concerned with universal qualities or a constant and habitual process, the present simple tense is in constant use.

In a text reconstruction tasks, however, learners listen to or read a text before they start writing. After the text is removed, the learners are supposed to reconstruct the text in their own words. "The underlying assumption is that in processing a text for meaning, learners have to store the propositional content but not the linguistic forms to encode the content. Required to reconstruct the text, therefore, they are forced to draw on their own linguistic resources" [8].

Second language writers have to challenge higher-level skills of planning and organizing, as well as lower level skills of spelling, punctuation, and word choice. One of the challenges in working on second language acquisition is to address the concurrent need for maintaining complexity, fluency, and accuracy in EFL learners' language. The desire to investigate motors of change contributes to the challenge.

The current study was, therefore, an attempt to shed more light on the notion of communicative tasks in EFL language teaching and assessment with a specific attention to using tasks in one of the most complex and critical language skills, i.e., writing. Moreover, the major objective in the current study was to investigate whether different types of assessment tasks would have any significant impacts on the fluency, accuracy, and complexity aspects of Iranian EFL learners' writing performance.

II. METHOD

The 65 subjects in this study were selected from both male and female population of Iranian EFL undergraduates (n=200), ranging from 19 to 36 years of age, who attended the course of advanced writing in 2010 at Alborz Institute of Higher Education, Qazvin, Iran

A modified Test of English as a Foreign Language (TOEFL) was initially administered to randomly sample the subjects. The 65 students whose TOEFL test scores dwelled in a continuum of one standard deviation below and above the sample mean score were selected as the experimental group in the study. The test contained 50 multiple choice items of listening comprehension, 40 multiple choice items on structure and written expressions, and 50 multiple choice items of reading comprehension.

Next, three tasks including, a Topic Writing Task (TWT), a Picture Description Task (PDT), and a Text Reconstruction Task (TRT) were designed and piloted with 18 students similar to the research subjects in their levels of language proficiency, gender and age. The participants' performance on these writing tasks was rated for degrees of accuracy, fluency and complexity by three raters who were professional in EFL language testing.

In order to measure the aspect of Accuracy in participants' writing performance, the traters counted the number of errors in every 100 words. This would indicate how learners were accurate in their writing performance.

For the aspect of Fluency, the raters counted the total number of words in the text and then divided them by the total number of seconds/minutes it took to be written. Finally, to measure the aspect of complexity, the T-unit was selected as the rating scale. The raters were supposed to count every clause which was attached to a T-unit (an independent clause). Next, the reliability estimates for the tasks scores proved the results desirably trustworthy and dependable.

Finally, an Analysis of Variances (ANOVA) was computed for three groups of writing task scores and three aspects of accuracy, fluency and complexity in the participants' writing performance. The obtained results demonstrated statistically significant differences in the participants' performance on different types of tasks.

III. RESULTS AND DISCUSSION

As already mentioned, a TOEFL test was administered to screen the participants based on their levels of language proficiency, and in order to randomly sample an experimental group in a continuum of one standard deviation below and

one standard deviation above the mean score.

Eventually, 65 subjects were selected to participate in this study. TABLE I: displays the descriptive statistics for the TOEFL scores.

TABLE I: DESCRIPTIVE STATISTICS FOR TOEFL SCORES

Test	M	Max	Min	SD	N
TOEFL Valid N (List-wise)	60.00	98.00	21.00	4.01	200

In TABLE I:, statistics for TOEFL scores included the mean (M), maximum scores (Max), minimum scores (Min), standard deviations (SD), and the number of EFL learners (N=200). The measures of mean, minimum, and maximum scores all revealed that the test scores were reliably well-centered while the value of standard deviation showed a rather wide range of TOEFL scores.

In order to determine the reliability of the TOEFL test results, Cronbach's Alpha test was conducted for the total number of test items. The obtained Cronbach α=.226, insignificant at p-value<0.05, demonstrated the relative inconsistency of the participants' performance on total test items.

Next, three Topic Writing Task (TWT), Picture Writing Task (PWT), and Text Reconstruction Task (TRT) were administered and rated. To estimate the reliability of the obtained results, two measures of inter-rater (Table II) and inter-item reliability (TABLE III) were conducted.

As TABLE II demonstrates, the adjusted Spearman-Brown reliability index of the raters' scores to the writing tasks was 0.98, statistically significant at a two-tailed p<0.05.

Similarly, in inter-item reliability computation for three groups of task scores, the average Item Total correlations of r=.78 indicated the individual writing task scores discriminated in a manner quite similar to the total scores on writing tasks, though they had done so to varying degrees with correlations ranging from r=.21 to r=.79 (TABLE III). It was also noticed that all the correlation coefficients were statistically significant at a two-tailed p<0.05.

In order to study the possible variations among the participants' performance on three types of tasks, the impacts of the three task types on aspects of accuracy, fluency and complexity were analyzed with an Analysis of Variances (ANOVA).

TABLE II:: INTER-RATER RELIABILITY ESTIMATE

Tasks	r ^{1and2}	r ^{1and3}	r ^{2and3}	Average r	Adjusted S
TWT	0.62	0.69	0.65	0.94	*0.98
PWT					
TRT					

TABLE III: INTER-ITEM RELIABILITY ESTIMATES

Tasks	M	Max	Min	SD	N	Skew	r $^{item-total}$
TWT	3.00	6.00	0.00	0.43	65	- 0.18	*0.78
PWT	2.00	6.00	1.00	0.23	65	- 0.23	*0.21
TRT	3.00	6.00	0.00	0.28	65	- 0.22	* 0.45

A. Task Types and Accuracy

TABLE IV: ACCURACY IN THREE TASK TYPES

Task	M	SD	N
TWT	1.15	1.485	65
PDT	0.48	0.981	65
TRT	0.39	0.803	65

TABLE V: AVOVA FOR ASPECT OF ACCURACY

Effects		Value	F	Hypothesis df	Error df	Sig	Noncent Parameter	Observed Power
accuracy	Pillai's Trace	.244	12.503	2.000	78.00	.000	25.127	.995
	Wilks' Lambda	.756	12.503	2.000	78.00	.000	25.127	.995
	Hotelling's Trace	.322	12.563	2.000	78.00	.000	25.127	.995
	Roy's Largest Root	.322	12.563	2.000	78.00	.000	25.127	.995

a Computed using alpha< 0.05

Demonstrated in TABLE IV, the participants' mean score in TWT was 1.15 which is much higher than the mean score for PDT (0.48) and TRT (0.39). Similar to mean scores, the measures of standard deviation for TWT (SD=1.48) was considerably higher than those for PDT (0.98) and (0.80) which showed a wide variety of performance on TWT.

Similarly, in TABLE V, all the F ratios computed in multivariate ANOVA were significant at p-value<0.05 which proved a vast range of variations in degrees of accuracy traced in the participants' performance on three types of tasks.

B. Task Types and Fluency

In TABLE VI, the comparison made for the impacts of different types of tasks on the fluency of the participants' writing performance revealed that Fluency mean score was radically high in PWT (27.32), comparing to the mean scores for TWT (16.23) and TRT (15.21). Regarding measures of standard deviation, the observed pattern followed those in the aspect of Accuracy, that is, a rise for PWT (SD=15.61) comparing to those for TWT (SD=5.82) and TRT (6.18).

As it was expected, TABLE VII demonstrates all measures of F ratios as significant at p-value<0.05 which proved a variation in Fluency aspect of participants' performance due to different task types.

TABLE VI: FLUENCY IN THREE TASK TYPES

Task	M	SD	N
TWT	16.23	5.82	65
PDT	27.32	15.61	65
TRT	15.21	6.18	65

TABLE VII: ANOVA FOR ASPECT OF FLUENCY

Effects		Value	F	Hypo thesis df	Error df	Sig.	ent. Para	ved Powe
fluency	Pillai's Trace	.400	25.957	2.000	78.00	.000	51.914	1.000
	Wilks' Lambda	.600	25.957	2.000	78.00	.000	51.914	1.000
	Hotelling's Trace	.666	25.957	2.000	78.00	.000	51.914	1.000
	Roy's Largest Root	.666	25.957	2.000	78.00	.000	51.914	1.000

a Computed using alpha < 0.05

C. Task Types and Complexity

Concerning the aspect of Complexity, demonstrated in TABLE VIII, the mean score for TWT (4.85) was higher than the mean for PWT (2.83) and for TRT (1.74). Measures of standard deviation showed the similar pattern of rise for TWT (SD=4.56) comparing to those for PWT (SD=2.82) and TRT (SD=1.49).

In TABLE IX, the measures of F ratios proved to be statically significant at p-value<0.05 which can be interpreted as a wide variation in degrees of Complexity traced in the participants' performance on three types of tasks.

To sum up, the findings in this study proved different types of writing tasks would have significant impacts on aspects of accuracy, fluency and complexity of Iranian EFL learners' writing performance. Evidence was in favor of a high accuracy and complexity when the participants were actively involved in Topic Writing tasks, and fluency inflation when the EFL learners were engaged in Picture Writing task.

TABLE VIII: COMPLEXITY OF THREE TASK TYPES

Task	M	SD	N
TWT	4.85	4.56	65
PDT	2.83	2.82	65
TRT	1.74	1.49	65

TABLE IX: ANOVA FOR ASPECT OF COMPLEXITY

Effect		Value	F	Hypot hesis	Error df	Sig.	Nonce nt. Param	Observ ed
compl	Pillai's Trace	.360	21.919	2.000	78.00	.000	43.837	1.000
	Wilks' Lambda	.640	21.919	2.000	78.00	.000	43.837	1.000
	Hotelling's Trace	.562	21.919	2.000	78.00	.000	43.837	1.000
	Roy's Largest Root	.562	21.919	2.000	78.00	.000	43.837	1.000

a Computed using alpha<0.05

IV. CONCLUSION

As Nunan [2] constantly urges, utilizing different types of task opens a window of opportunity for EFL learners to have real-life interactions. Based on the experimental findings in this study, EFL learners experience enjoyable difference to perform differently on tasks of different nature. They welcome new writing tasks more than traditional writing didactics. It is, therefore, recommended that EFL teachers make use of a variety of writing tasks in their classroom as far as they can.

If their purpose is to enhance learners' fluency, for example, a picture writing task would be an appropriate activity to join in the classroom; where their emphasis is on accuracy, as it is mostly the case, a topic writing task is the best choice for effortlessly eliciting the required data. Finally, when EFL teachers desperately struggle to redirect their students' attention to a particular grammatical point, a topic writing task can work remarkably.

Skehan and Foster [9] offer suggestions for balancing the goals of accuracy, complexity, and fluency in EFL contexts. They believe that it is hard to make balance between these three elements, because of different task demands, as for example, one task may put more emphasis on structure to enhance accuracy, while others pay more attention to time in order to foster fluency. It is therefore the teachers' responsibility to implement a task wherein there exists a sort of balance and appropriate ground to improve these three elements simultaneously and equally.

Hyland [10] lists a variety of techniques to provide teachers' feedback to EFL learners. The most common ones are commentary, rubrics, and minimal marking. Commentary is probably the most common type of teacher feedback where a handwritten commentary is given on students' paper itself. This kind of feedback is best seen as responding to students' work rather than evaluating what they have done, stating how the text appears to teacher as reader, how successful teacher thinks the text has been written, and how it could be improved [10].

Rubrics are defined as a variation of commentary and often accompanying it on the final draft. Rubrics provide the learners with the agreed upon criteria that have been used to assess the assignment. Different rubrics can be used in different writing genres. While they restrict the range of issues that can be addressed, they are useful in making explicit decisions and showing what the EFL teachers' values are in a particular writing task [10].

Minimal Marking refers to the teacher's in-text, form-based feedback. Minimal marking has mostly followed the research line which suggests that indicating the location and type of errors, rather than direct correction, is more effective in stimulating learners' response and also in developing their self-editing strategies. One way of accomplishing this is to use a set of simple "correction codes" which make correction neater and less threatening than masses of red ink and help students to find and identify their mistakes [10].

REFERENCES

[1] P. Skehan, *A Cognitive Approach to Language Learning.* Oxford: Oxford University Press, 1998a.
[2] D. Nunan, *Syllabus Design.* Oxford: Oxford University Press, 1998.

[3]   R. Ellis, R. *Task-based language learning and teaching.* Oxford: Oxford University Press, 2003.

[4]   W. Renandya and J. C. Richards, *Methodology in Language Teaching.* Cambridge: Cambridge University Press, 2002.

[5]   J. Murphy, "Task-based learning: The interaction between tasks and learners," *ELT Journal*, 58 (4) 36-49, 2003.

[6]   S. Gass, "Interaction perspectives in second language acquisition," In R. Kaplan (Ed.), *Handbook of applied linguistics*, Oxford: Oxford University Press. 2002

[7]   H. Farhady, "Language assessment: A Linguametric perspective,"

Language Assessment Quarterly, vol. 2 (2), pp. 38-45, 2005.

[8]   P. Lightbown, "Classroom SLA research and language Teaching," *Applied Linguistics*, vol. 21 (4), pp.431-50, 2000.

[9]   P. Skehan and P. Foster, "The influence of task structure and processing conditions on narrative retellings," *Language Learning*, vol.49 (1), pp. 98-101, 1999.

[10]   K. Hyland, and F. Hyland, "Feedback on L2 writing: Contexts and issues - an introduction" in K. Hyland, and F. Hyland (Eds), pp.1-19, 2006.