Integrating Divergent Assessment (DA) and Task Based Language Teaching (TBLT) in Argumentative Essay Writing Classrooms: An Experimental Study to Develop Iraqi EFL Learners' Writing Skills

Esmaeil Bagheridoust[1] & Hossam Al-Sultani[1]

[1] Department of TEFL, College of Foreign Languages and Literature, Islamic Azad University, South Tehran Branch, Iran

Correspondence: Hossam Al-Sultani, Department of TEFL, College of Foreign Languages and Literature, Islamic Azad University, South Tehran Branch, Iran. E-mail: qwsksjdjdldk@gmail.com

Abstract

The researchers delved into the practice of divergent assessment (DA) and participation in assessment while emphasizing task-based language learning (TBLT) as instruction in argumentative essay courses in an Iraqi EFL context. In doing so, they formed 3 classes, namely a traditional class, a DA exercise with a traditional class, and a TBLT exercise with a traditional class, to investigate the differences that such practices can induce in the writing quality of EFL learners. performance Post-test analysis and one-way ANOVA show that the CONV TBLT class (M = 3.97) significantly outperformed the CONV DA group (M = 2.98) in terms of better argumentative writing. This study shows how EFL learners respond to the integration of DA and TBLT with compliant writing instruction. The researchers used two parallel argumentative writing tasks before and after the DA and TBLT treatment. The results of the study rejected the researchers' null hypotheses, showing that both DA and TBLT were reliably effective in increasing Iraqi EFL students' argumentative essay excellence, with TBLT having a slightly higher neck than DA. The final result of the study clearly shows that the CONV TBLT group showed a significantly higher mean score than the CONV DA and CONV groups. The CONV DA group showed a significantly higher mean score than the CONV group.

Keywords: divergent assessment (DA), task-based language instruction (TBLT), argumentative essay writing, conventional writing, intermediate EFL learners

1. Introduction

1.1 Statement the Problem

The high rate of essay writing applicants of different standardized tests; namely, IELTS, TOEFL, and GRE in particular, in non-English speaking countries, Iraq as the locus of this study, and the low rate of efficient and effective writing teachers are creating anxiety, lack of self-confidence, and poor performance in writing of Essay writing in time of real test. This major problem of the EFL system in Iraq has motivated a number of researchers to embark on investigating the reasons behind the poor performance of both Essay writing teachers and Essay writing candidates of different tests during the course semesters and finally the real exam time.

This study mainly focused on discovering and then applying the pragmatic strategies of argumentative essay writing based on the tasks delivered in classes, especially TBLT through using the kernel of divergent assessment to give the essay writing candidates both the skills and ability in order to compose comprehensible, clear, and undeviating essays. The techniques used in this regard provide the candidate with an easy-to-use format and strategy for structuring paragraphs that express the main points of the argument in a clear way. The controversial writing genre is ubiquitous and is commonly encountered by English learners in both secondary and higher education educational settings. Indeed, it is safe to assume that very few students studying English are spared the task of writing argumentative essays at some point in their academic career. One of the potential attractions of essays for teachers at both levels is their role in connecting the classroom to ideas, problems, and events occurring in the world outside academia. More than two decades ago, Crowhurst (1988) observed that "no style of writing offers more opportunities to write about real issues for real audiences than arguments" (p. 35). The belief that influential writing brings about success in both academic situations and real life (Crowhurst, 1988) is

what numerous EFL instructors would be glad about. Some scholars in the filed have correspondingly stated that the "argumentative essay is one of the most common forms of the curriculum genre that undergraduate students write" (p. 330). Moreover, among this handful of studies, most of them have examined the effectiveness of DA in its generality on the production of speech acts (e.g. Moradian, Asadi, & Azadbakht, 2019; Tajeddin & Tayebipour, 2012) and rarely on the comprehension of various speech acts or implicatures in Iranian EFL context. But even among learners whose mother tongue is English, argumentative writing can be inspiring, and it is conceivably not astonishing that students in various levels and nations have performed off-color in this genre (Crowhurst, 1988). Yet, despite the prominence of the genre, there is amazingly little TESOL research on what precisely the argumentative essay is; how is written; what it requires; and most importantly why the archetypal argumentation structure is so prevalent.

Despite the lack of research in this connection, there are numerous resources for language instructors like textbooks, that provide guides and drills in argumentative writing. However, this research manily questions how exactly the argumentative essay training reflects the type of writing that are composed outside of academic writing instruction platforms reflect real-world context. Nevertheless, providing a sketch that incorporates the entire argumentative essay is challenging. Like Crowhurst (1988), the researcher in this article is not interested in depicting a discrepancy between the terms convincing and argumentative. Although they are often used interchangeably, both are often overloaded and designate writings that surpass disciplines, contexts and addressees. Thus, an argumentative essay is generally defined here, as defined by Hyland (1990), which is merely an essay, the purpose of which is to persuade the reader of a centralized opinion. The focus of this study is the construction of the argumentative essay so often taught in academic writing contexts for EFL and ESL learners worldwide and how faithfully it brings into line with authentic real-world settings.

1.2 Significance of the Problem

This inquiry is estimated to have theoretical and practical standing in DA and TBLT, as well as learning and teaching argumentative essays. It is imperative that we get an intuition into the nature of DA and TBLT and the corresponding benchmark in rating EFL learners' ability to write English essays as compared to the near-native writers of English. This probe, is of crucial significance to language educators and EFL instructors not only in Iraq but in different EFL hubs where teaching and learning argumentative essay writing is challenging for both EFL teachers and learners. The results and findings of this study will be first and foremost to the benefit of EFL learners and teachers, EFL language schools, EFL colleges, companies, and organizations where argumentative essay writing is a serious issue. The use of DA and TBLT in argumentative essay writing classes would definitely promote professionalism.

1.3 Related Literature Review

Hudson was the first to develop convergence and divergence estimators in 1966. As a student, Hudson had previously struggled with some tests; therefore, he began to study the importance and type of test. Convergent and divergent assessments are named after the two ways of thinking a student needs to complete the assessment. It can be said in advance that if there is only one correct answer in the assessment, which is clearly and conventionally accepted information for the student, it is called a convergent assessment. On the other hand, if the evaluation uses, for example, an inverse approach, there are several ways to answer correctly; this is called divergent evaluation. Convergent and divergent evaluation both took a back seat to convergent and divergent thinking. We would like to discuss some related studies to cast more light on the nature of these two kinds of assessment and their interaction with test taker's personality factors and test designer's contribution to assessment process. In a previous study, Tzeng, G. and Huang, J. (2011) have explored the relationship between learners' motivation and learning strategies and the type of assessment they are taking. She focused on convergent and divergent assessment in her studies. She described in her article, the differences convergent and divergent assessments in-depth. According to her, convergent and divergent assessments are described as "traditional test" for the former and "group performance assessment" for the latter (p. 251). Huang explained that assessment offers a renewed interest in the fact that teaching, learning, and assessment could be combined together. She emphasized that researchers have focused on the teachers' roles in classroom rather than understanding of students' interpretations and reactions to an assessment which is as much important in the process of teaching and learning. Later in her studies, she expresses that" without this understanding of the role of the learners, teachers are very likely to interpret student performance inaccurately" (p.253).

The results of the study show that the learners' motivation and strategies were higher in speaking and lower in listening when students experienced divergent assessment. On the other hand, she emphasizes the learner's understanding rather than the agenda of the assessor. Here, the important thing is to discover what the learner

knows, understands and can do. It is characterized by less detailed planning, where open questioning and tasks are of more relevance. The implications of divergent teacher assessment are that a constructivist view of learning is adopted, with an intention to teach in the zone of proximal development. As a result, assessment is seen as accomplished jointly by the teacher and the student, and oriented more to future development rather than measurement of past or current achievement" (Torrance & Pryor, 2001, p. 617). Huang's explanations about the description of the assessment are rather vague and do not allow determining whether (p. 256). Divergent assessment used in his research follows the concept of Vygotsky (1978, p. 86) of the Zone of Proximal Development (ZPD). Instead of divergent evaluation, it would be more correct to talk about learning-promoting evaluation, because the latter type refers to dynamic evaluation, which is completely different from learning-promoting evaluation in its nature (Poehner and Lantolf, 2005, p. 260). Huang (2011) researched and investigated two types of classroom assessment, viz. convergent and divergent evaluation, in a correlational study. He conducted a study to determine if DA and CA affect students differently in terms of motivation and learning strategies. He recruited 105 university freshmen from Taiwan who were taught listening and speaking by the same teacher. In these unaffected groups, he allowed each student to undergo both types of assessment, namely the more traditional test (CA) and group performance assessment (DA). He will report his motivations and strategies immediately after both evaluation events. Because the results of his study showed that students with high self-efficacy scored better on CA and those with low scores on DA. Publishing the results of a self-efficacy study, he found that students responded differently in listening and speaking lessons. Specifically, he reported that in speaking, students' motivation and strategy were higher for DA than CA. This result was completely opposite to the listening.

Task-based language learning (TBLT) has received mixed reviews about whether it is really effective in upholding students' communication skills. Studies on the effectiveness of task-based language learning (TBLT) and communicative language learning (CLT) for Japanese learners of English in Japan has produced amazing results. Both approaches were found to be incompatible with English classes taught in Japan, as they focused on students' mastery of grammar. Their goal was to produce grammatically correct sentences. The study successfully concluded that the most compatible approach for the classrooms is the traditional form-centered approach such as Grammar Translation Method (GTM), Performance Practice Production (PPP) and Test Teaching Test (TTT). This study was reliable with a previous study that found TBLT unclear for Japanese students. Furthermore, Japanese L2L researchers criticize TBLT for its unsupportive goal-related task. Their conclusions seem to summarize that: (1) TBLT has not been accurately employed in the English as a Foreign Language (EFL) classroom because students do not have an urgent need to use English; (2) TBLT did not meet students' learning style and expectation; (3) TBLT lowered the motivation of Japanese students who focused on preparing for the English language test; (4) TBLT was compatible with Confucian heritage, believing that lectures or teachers always have more authority than students; (5) TBLT has a grammar-centered learning model; (6) TBLT violates the principle of grammar curriculum learning; and (7) TBLT use may lead to language fossilization rather than language acquisition. The Performance Practice Production (PPP) approach was found to be ineffective in improving students' communication skills, and TBLT was very successful in second language learning and classroom-based research. TBLT has been shown to have some improvements: (1) CLT supports this approach well; (2) the approach was a reaction to the failure of the PPP and TTT approaches; (3) TBLT treats the target language as a means of communication rather than as an object of learning; (4) TBLT does not control students in presentations and exercises, where students receive a grammar lesson better than the PPP approach. Under the impact of TBLT integrated with DA, the writing task is seen as a decontextualized form; it ignores context and audience and emphasizes the learner's thesis rather than how it is done, so the importance of a process-based approach must also be emphasized in this regard. An argumentative essay is a very important part of standardized tests like the TOEFL. This is one of two essays written in the essay section. However, this is not just another writing question. An argumentative essay is different from the regular essays you write and the essay asked in the same part of the test. Therefore, it is very important to understand the needs of this essay before attempting it.

1.4 Research Questions and Formulated Hypotheses

Based on the tenets of the study, these research questions are challenging:

Does the integration of Divergent Assessment with Conventional Argumentative essay writing class significantly improve the quality of Argumentative Essay writing task?

Does the integration of TBLT with Conventional Argumentative essay writing class significantly improve the quality of Argumentative Essay writing task?

Is there any significant difference between learner's argumentative essay writing performance in conventional class vs. conventional class plus divergent assessment?

Is there any significant difference between learner's argumentative essay writing performance in conventional class vs. conventional class plus TBLT?

Is there any significant difference between learner's argumentative essay writing performance in conventional class plus Divergent Assessment vs. conventional class plus TBLT?

The researcher conceived the following null-hypotheses accordingly:

The integration of DA with Conventional Argumentative essay writing class DOES NOT significantly improve the quality of Argumentative Essay writing task?

The integration of TBLT with Conventional Argumentative essay writing class DOES NOT significantly improve the quality of Argumentative Essay writing task?

There is NO significant difference between learner's argumentative essay writing performance in CONV class vs. CONV + DA?

There is NO significant difference between learner's argumentative essay writing performance in CONV class vs. CONV + TBLT?

There is NO significant difference between learner's argumentative essay writing performance in CONV class + DA vs. CONV class + TBLT?

2. Method

2.1 Participants' Particulars

In the first frame of the experiment, the researcher invited almost one hundred (100) Iraqi EFL learners of 3rd and 4th grade from an Iraqi high school called Imam Al-Kadhim College - Babylon to participate in the first experiment in which homogeneous was chosen sample population. The researcher selected eighty-seven (87) participants out of one hundred students who were the final recruited participants as the main students of the study. Candidates were selected based on the combined score of the language proficiency test in the first cell of the exam. The researcher intended to compare their experiences with each type of writing task and at the same time measure the common difference between the two scores obtained by the same candidates in a TBLT class integrated into a traditional writing class compared to a deviant assessment integrated into a class. traditional essay class. essay class compared to a traditional essay class. An argumentative writing course. The researcher asked about the language background and future plans of the candidates. Based on the purposeful sampling, the researchers formed two intact groups to partake in the experimentation. Due to the restriction of gender factor in Iraq, only male EFL learners were nominated. The average age of the participants was between 18 and 25. Some participants were eliminated after the homogeneity test was run and as mentioned before 87 homogenized participants based on their language proficiency level were finally chosen as the core subjects of the study. In the next phase, the researchers divided three treatment (experimental) and control (normal) groups, each with 28, 29, and 30 students. One experienced EFL instructor was invited to lead and manage three classes namely CONVTIONAL, CONV DA, and CONV TBLT groups. The sample population is truly representative of the target population because the researcher conducted the homogenization test and the normality test. The result of the homogenization confirmed the degree of generalizability necessary to confirm the results of the study. Appropriate identification of research participants is critical to the science and practice of psychology, particularly for generalizing the findings, making comparisons across replications, and using the evidence in research syntheses and secondary data analyses. If humans participated in the study, report the eligibility and exclusion criteria, including any restrictions based on demographic characteristics.

2.2 Sampling Procedures and Data Collection Instruments

To obtain measurable data, the researcher used availed from TOEFL independent argumentative essay writing tasks to measure participants' current level of writing and achievement level before and after instruction. The standard TOEFL Independent Argumentative Essay test was given to Iraqi EFL learners as a pretest. This is one of two essays written in the essay section. However, this is not just another writing question. The question in an argumentative essay is in the form of a statement written by the author. It represents the author's position on the matter. Testers must discuss the credibility of the claim. This means they need to discuss the logic behind the claim. It also includes the writer's thought process, which they must recognize. Either the writer is right to keep his thoughts or he is wrong and misses some important things which can make the argument baseless in some cases. The tester can discuss areas where the argument is not valid and suggest an alternative argument or add

something more logical to the same argument (Mitchel 2015).   And finally, since the study lasted almost three months (one semester), the researcher used a parallel TOEFL Independent argumentative test to ensure that the teaching of the experimental and control groups was somewhat effective. This second writing test was used as a post-test to measure the difference between the new and conventional processing.

2.2.1 TOEFL Independent Argument Essay Scoring Rubrics

In the TOEFL Independent Argumentative Essay writing section, the candidate has a defined writing task to write. Evaluators give candidates the score between 1 to 5. This score is converted to a scale of 0-30 when taking the TOEFL test. This is a raw TOEFL score and this raw TOEFL score is converted to a scale score on a scale of 0-30. Candidates can check the TOEFL iBT Writing Conversion Chart. (See Appendix)

2.3 Design of the Study

The investigator directed a quasi-experimental research based on a pre-post-test survey. As mentioned before, there were two intact classes of participants. Random purposive sampling was used to compare the learning outcomes of students exposed to a "regular" curriculum with students exposed to a new teaching approach integrated with DA, TBLT, and conventional argumentative essay writing techniques in the same syllabus. The following step-by-step procedure reflects the structure of the research in a hierarchical way:

Pre-test Treatment Post-test
O1 XT O2 Experimental Group
Randomly assigned to
Control Group
O1 XT O2
Pre-test Treatment Post-test
Pre-Posttest EG/CG design

(Adapted. from Johnson & Christensen, 2004, p. 283)

Where: O1 and O2 represent the pre-test and post-test assessment of the dependent variable.

XT represents the treatment condition. XC represents the control or standard treatment condition.

2.4 Stepwise Procedure

As variables determine the constraints of this study, the researcher went through the following step-by-step procedure to conduct the experimentation. These phases include the pretest, treatment, and posttest of the study.

(1) The study was conducted in a 12-week language school semester program.

(2) The researcher selected 87 participants and formed 3 categories, viz. traditional writing, a mixture of different assessments and a traditional writing category, and finally a mixture of TBLT and a traditional writing category:

Class A: CONV Group

Class B: CONV + DA Group

Class C: CONV + TBLT Group

(1) The three groups were given a pre-test covering the TOEFL topic of planning and writing an independent argumentative essay without knowing how to complete the writing task. All of these were later evaluated as a starting point for statistical comparison.

(2) All participants were taught and introduced to argumentative essay writing and reasoning using the TOEFL Independent Argument essay type, where they had to critically read an argument and analyze the chain of reasoning used in the argument.

(3) Treatment lasted 12 sessions over 12 weeks and was carried out individually and in a third group both individually and collaboratively. It lasted 30-40 minutes in each session. The procedure involved the following steps:

(4) In the Traditional Argumentative Essay Writing Approach group (Class CONV), participants received instructions on how to complete the task in writing. They were asked to analyze a TOEFL Independent argumentative writing sample based on principles previously taught. In this group, the teacher did not interfere with the participants' writing or give suggestions to complete the task.

(5) In the CONV DA integrated class, the researcher integrates deviant assessment with the traditional TOEFL independent argumentative essay writing course to carry out various classroom activities to promote the development of language use so that writers can take an active and dynamic approach to their writing. task instead of one shot.

(6) In the CONV TBLT integrated class, the researcher integrates the Task-Based Language Teaching program into the traditional TOEFL independent argumentative writing class, which is able to implement various language development activities in the classroom so that the writers can participate in their writing task an active and dynamic approach rather than a single shot.

(7) In the first integrated group, or CONV DA, participants completed a writing task in four stages while receiving feedback via deviance assessment. Another integrated class, CONV TBLT, used writing assignments to generate more ideas and thoughts for writing by setting the stage for an argumentative essay topic and providing guidance for writing, editing, revising, and editing (Sapiens 2014).

Prewriting is the first step in the writing process, usually followed by editing, revising, and editing. In this stage of the writing, participants divided information and ideas into sentences and paragraphs. During outlining, participants form their ideas into complete thoughts, such as sentences and paragraphs. The teacher then provided feedback on students' writing to support students' writing development and increase their confidence as writers. In the third, review stage, they reread their writing sample and the written feedback given by the teacher, and finally made some changes (in content, organization, sentence structure and word choice) to improve it. And finally, in editing, participants try to improve the draft by correcting mistakes and making words and sentences more clear, precise and effective. According to Meyers (2010), "Some of the most effective edits are to tighten and shorten a piece to make it better" (p. 87). After the treatment period, all groups received a post-test: TOEFL Independent argumentative essay parallel writing test. At the end, the results of both the pre- and post-test were compared according to the effectiveness of the program.

3. Results

In the Results section, summarize the collected data and the analysis performed on those data relevant to the discourse that is to follow. Report the data in sufficient detail to justify your conclusions. Mention all relevant results, including those that run counter to expectation; be sure to include small effect sizes (or statistically nonsignificant findings) when theory predicts large (or statistically significant) ones. Do not hide uncomfortable results by omission. Do not include individual scores or raw data with the exception, for example, of single-case designs or illustrative examples. In the spirit of data sharing (encouraged by APA and other professional associations and sometimes required by funding agencies), raw data, including study characteristics and individual effect sizes used in a meta -analysis, can be made available on supplemental online archives.

3.1 Recruitment

Researchers recruited the participants of this study from the limited population of an Iraqi high school called Imam Al-Kadhim College – Babylon for a limited period of December 2022 thru February 2023.

3.2 Testing Assumptions (Statistics and Data Analysis)

This study mainly investigates the divergent assessment (DA) and task-based language learning (TBLT) approach on the quality of Iraqi EFL students' argumentative essay writing. Three groups of participants participated in this study (traditional, integration of dynamic assessment with traditional, and integration of TBLT with traditional). They took a pretest and a posttest on argumentative writing tasks. The research questions presented were analyzed using a one-way ANOVA parametric test with two main assumptions; normality of data and homogeneity of group variances. The normality of the data was indicated by the ratios of the skewness and kurtosis indices to their standard errors. As shown in Table 1, the ratios vary from /- 1.96. So it can be concluded that the current data is normally distributed.

3.3 Ancillary Data Analyses

Table 1. Descriptive Statistics; Testing Normality Assumptions

Group		N	Skewness			Kurtosis		
		Statistic	Statistic	Std. Error	Ratio	Statistic	Std. Error	Ratio
Conventional	Pretest	28	.426	.411	0.84	-.821	.808	-0.75
	Post-test	28	-.074	.401	-0.29	-.886	.758	-1.33
CONV + DA	Pretest	29	.547	.414	1.14	.082	.745	0.17
	Post-test	29	.413	.474	1.12	.286	.645	0.21
CONV + TBLT	Pretest	30	.662	.407	1.08	.065	.733	0.18
	Post-test	30	-.747	.497	-1.92	1.201	.933	1.33

3.4 Pretest of Argumentative Essay Writing

To demonstrate this, a one-way analysis of variance was performed to compare the conventional, dynamic assessment integration with the conventional and the TBLT integration with the conventional (CONVENTIONAL, CONV DA, CONV TBLT) groups. they had the same writing skills before the study and treatment. Before discussing the results, it should be mentioned that the assumption of homogeneity of variances was met (Levene's F (2, 84) = .55, p> .05) (Table 2).

Table 2. Levene's Test of Equality of Error Variances

F	df1	df2	Sig.
.612	2	84	.634

Table 3 shows the descriptive statistics for the three classes on the pretest. The conventional class (M = 2.88, SD = 1.03) showed the highest mean on the pretest of essay writing. This was followed by CONV + DA (M = 2.78, SD = .90) and CONV + TBLT (M = 2.75, SD = .90) groups.

Table 3. Descriptive Statistics, Pretest of Argumentative Essay Writing by Groups

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Min	Max
					Lower Bound	Upper Bound		
CONV	28	4.88	1.01	.175	2.27	4.18	2	5
CONV + DA	29	1.18	.912	.148	2.49	4.02	2	5
CONV + TBLT	30	1.75	.897	.116	2.21	4.09	2	5
Total	87	1.80	.915	.111	2.40	4.020	2	5

From the results of one-way ANOVA (F (2, 84) = 0.13, p> 0.05, ω2 = 0.02, which denotes a weak effect size) (Table 4), it can be concluded that there were not significant differences between the three classes when writing an essay. Therefore, it can be concluded that they had the same level of reasoning ability before the main study started. Figure 1, graphically depicts the same fact.

Table 4. One-Way ANOVA, Pretest of Argumentative essay Writing by Groups

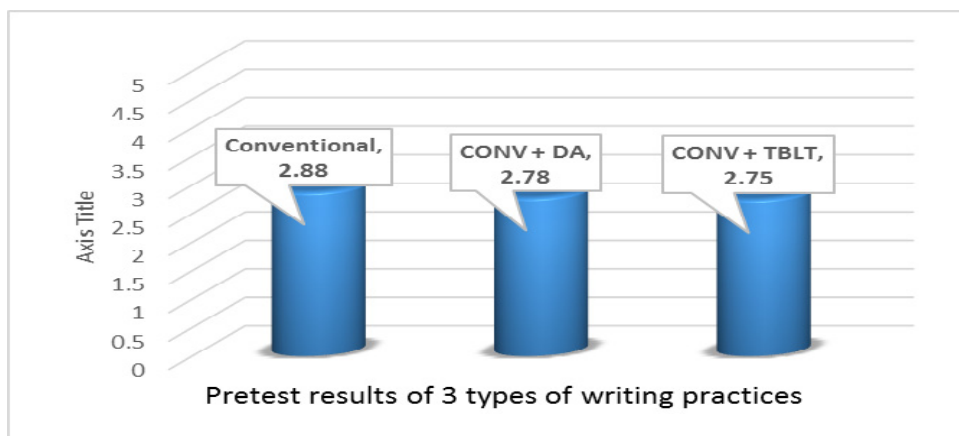|	Sum of Squares	Df	Mean Square	F	Sig.
Between Groups	.278	2	.155	.170	.931
Within Groups	74.381	84	.789		
Total	74.736	86			

Figure 1. Pretest of Argumentative Essay Writing by Groups

3.5 Testing Null-Hypotheses

A one-way analysis of variances was run to compare the process, product and process + DA groups' means on the post-test of argumentative essay writing in order to probe the two research questions posed in this study. Before discussing the results, it should be mentioned that the assumption of homogeneity of variances was met (Levene's F (2, 84) = .39, p > .05) (Table 5).

Table 5. Levene's Test of Equality of Error Variances

F	df1	df2	Sig.
.417	2	84	.774

Table 6 displays the descriptive statistics for the three groups on the post-test of argumentative essay writing. The CONV + TBLT group (M = 4.17, SD = 1.13) showed the highest mean on the post-test of argumentative essay writing. This was followed by CONV + DA (M = 2.98, SD = .80) and CONV (M = 2.29, SD = .87) groups.

Table 6. Descriptive Statistics, Post-test of Argumentative Essay Writing by Groups

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean Lower Bound	Upper Bound	Min	Max
CONV	28	2.29	.876	.166	1.95	2.63	1	4
CONV +DA	29	2.98	.807	.150	2.68	3.29	2	5
CONV + TBLT	30	3.97	1.033	.189	3.58	4.35	1	6
Total	87	3.10	1.138	.122	2.86	3.34	1	6

The results of the one-way ANOVA (F (2, 84) = 24.92, p< .05, ω2 = .35 representing a large effect size) (Table 7) it can be concluded that there were significant differences between the means of the three groups on the post-test of argumentative essay writing.

Table 7. One-Way ANOVA, Post-test of Argumentative Essay Writing by Groups

	Sum of Squares	Df	Mean Square	F	Sig.
Between Groups	41.497	2	20.749	21.18	.000
Within Groups	69.922	84	.832		
Total	111.420	86			

Although an F value of 21.18 indicated significant differences between the three classes means on the Essay Writing post-test, post-hoc Scheffe tests should be conducted to pairwise compare the classes to clarify the research questions. Based on the results presented in Table 8, it can be concluded that;

A: The CONV TBLT group (M = 3.97) significantly outperformed the CONV DA class (M = 2.98) on the essay writing posttest (MD = 0.98, p value; 0.05).

B: The CONV TBLT group (M = 3.97) significantly outperformed the CONV class (M = 2.29) on the essay writing posttest (MD = 1.68, p value; .05).

Based on these results, it can be concluded that null hypotheses 1 and 2, that the integration of CONV DA and CON TBLT in the CONV essay writing class did not significantly improve the quality of the argumentative writing task, were rejected as the post hoc analysis was rejected. evidence indicates a significant and reliable difference in student performance in the pretest and posttest. The CONV TBLT class significantly outperformed the CONV DA and CONV classes on the essay posttest. This is evidence to reject null hypotheses 4 and 5.

Table 8. Multiple Comparisons; Post-Hoc Scheffe's Tests

(I) Group	(J) Group	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
CONV + TBLT	CONV	1.681*	.240	.000	1.08	2.28
	CONV + DA	.984*	.238	.000	.39	1.58
CONV + DA	CONV	.697*	.242	.019	.09	1.30

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

C: The CONV + DA group (M = 2.98) significantly outperformed the CONV (M = 2.29) group on the post-test of argumentative essay writing (MD = .69, p < .05). Such a result rejects the 3rd null- hypothesis.
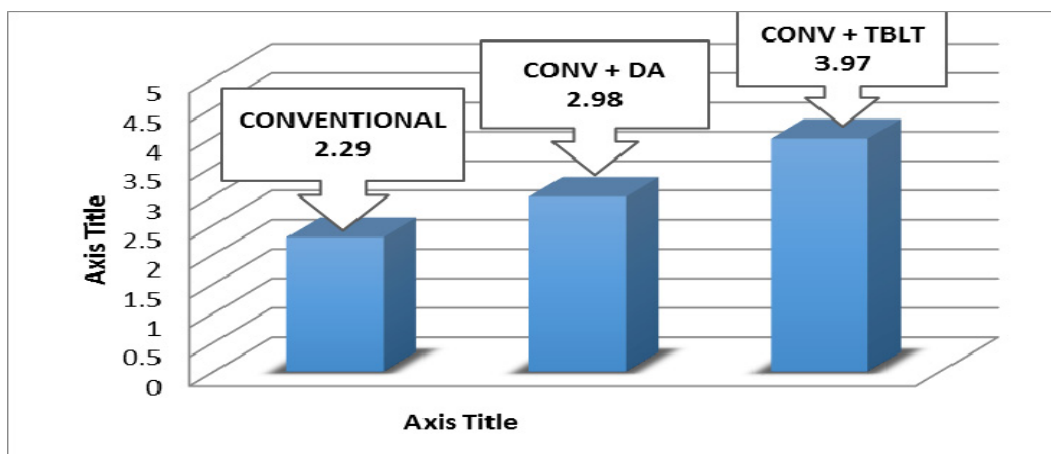

Figure 2. Post-test of Argumentative Essay Writing by Groups

The ultimate result is that the CONV + TBLT group significantly showed a higher mean than the CONV + DA and CONV groups. The CONV + DA group significantly showed a higher mean than the CONV group.

3.6 Inter-Rater Reliability

The researcher invited wo raters to rate the participants' essays on the pretest and post-test phases. The Pearson-Product Moment Correlation Coefficient was run to probe the inter-rater reliability indices for their ratings.

3.6.1 Pretest of Argumentative Essay Writing

There was a significant agreement between the two raters who rated the subjects' argumentative essay writings on the pretest (r (85) = .81, p < .05 representing a large effect size.

Table 9. Inter-Rater Reliability; Pretest of Argumentative Essay Writing

		Pretest Rater 2
Pretest Rater 1	Pearson Correlation	.877**
	Sig. (2-tailed)	.000
	N	86

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

3.6.2 Post-test of Argumentative Essay Writing

There was a significant agreement between the two raters who rated the subjects' argumentative essay writings on the post-test (r (.87) = .88, p < .05 representing a large effect size.

Table 10. Inter-Rater Reliability; Post-test of Argumentative Essay Writing

		Post-test Rater 2
	Pearson Correlation	.897**
Post-test Rater 1	Sig. (2-tailed)	.000
	N	86

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

3.7 Construct Validity

A factor analysis was run to probe the underlying constructs of the pretest and post-test of argumentative essay writing. The SPSS extracted only one factor which accounted for 72.56 percent of the total variance.

Table 11. Total Variance Explained

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	1.351	72.567	72.567	1.551	75.567	72.567
2	.649	27.433	100.000			

Table 12. Component Matrix

	Component
	1
Post-test	.78
Pretest	.79

As shown in Table 12, the Argumentative Essay Writing pretest and posttest loaded on one factor. These results indicated that both tests measured the same trait, implying their construct validity.

4. Discussion

The current research is a two-part research that includes both macro studies on teaching professionalism and micro studies on learning development. To support our claim, the researchers experimented with research questions and tests, followed situational statistical procedures and arrived at comprehensive statistical results. In recent years, language skills, especially writing, have been widely researched, integrating a number of teaching strategies and assessment practices.

Strategy training is based on the assumption that learning success depends primarily on appropriate strategy use and that unsuccessful students can improve their learning if they are trained to use effective strategies (Dansereau, 1985; Weinstein and Underwood, 1985).

As mentioned earlier, this study investigated the effects of integrating TBLT and DA in argumentative essay programs of Iraqi EFL students. The integration of Divergent Assessment and TBLT into traditional argumentative writing classes resulted in improved performance on writing assignments compared to traditional writing classes. Based on an in-depth analysis, the researcher made the following conclusions and findings:

(1) The use of Divergent Assessment (DA) as a sort of classroom practice and drill significantly changes the quality of argumentative essay writing of Iraqi EFL learners.

(2) The use of Task-Based Language Teaching (TBLT) as a kind of classroom practice and drill significantly and reliably changes the quality of argumentative essay writing of Iraqi EFL learners.

(3) Based on the posttest analysis by employing One-way ANOVA, the CONV + TBLT class (M = 3.97) significantly outperformed the CONV + DA (M = 2.98) group in producing higher quality argumentative essay writing tasks.

(4) In the same line of comparison, the CONV + TBLT group (M = 3.97) significantly outperformed the CONV (M = 2.29) group on the post-test of argumentative essay writing.

(5) The results obtained clearly indicate that the 1st, and 2nd null-hypotheses reflecting the integration of CONV + DA and CON + TBLT against the MERE CONV Essay Writing classes did not significantly improve the quality of Argumentative essay writing task were rejected, since the post hoc analysis indices indicate significant and reliable difference in the performance of learners in pretest and post-test.

(6) The CONV + TBLT group significantly outperformed the CONV + DA and CONV groups on the post-test of argumentative essay writing. This is the evidence for the rejection of 4th and 5th null-hypotheses.

(7) The final result of the study indicates clearly that the CONV + TBLT group significantly showed a higher mean than the CONV + DA and CONV groups. The CONV + DA group significantly showed a higher mean than the CONV group.

5. Conclusion and Findings

In the first part of the study, the researchers considered practicing both DA and TBLT, one based on the nature of assessment and the other on teaching, and regular learning tools for teaching argumentative essay tasks. In the Iraqi EFL context. The results of the study show that using both DA and TBLT has a great impact on teachers and students to engage in writing tasks with more confidence. In addition, the researcher found that students' enthusiastic participation in argumentative writing tasks due to TBLT and DA practices significantly improved, which is clearly shown in the post-treatment scores. People who will definitely benefit from the results of the study are EFL teachers, curricula, language schools and institutions and EFL learners in Iraq. In the following, I, the researcher, would like to mention some consequences for the mentioned people. Teachers of English in EFL contexts, especially in Iraqi contexts, need special care regarding assessment and teaching practices that encourage students to perform better and more appropriately in argumentative essay tasks. To implement DA and TBLT systems and tactics, teacher education program centers and language schools and institutions need to familiarize EFL teachers with the concepts of assessment and task-based teaching. Iraqi EFL learners should be given the opportunity to be more dynamic and enthusiastic in dealing with argumentative writing tasks under the influence and practice of DA and TBLT. The researcher asked the students to interact in a real class before completing the written tasks and try to get each member to participate in the task. Iraqi EFL students must actively and critically analyze the written questions for each other and then proceed to write the thesis.

References

Anderson, J. R. (1993). *Rules of the mind*. Hillsdale, NJ: Lawrence Erlbaum Associations. https://doi.org/10.4324/9781315806938

Bachman, L. F., & Palmer, A. S. (2010). *Language assessment in practice: Developing language tests and justifying their use in the real world*. Oxford: Oxford University Press. https://doi.org/10.4236/ojml.2013.31004

Crowhurst, M. (1988). *Research review: Patterns of development in writing persuasive/argumentative discourse*. Vancouver, British Columbia, Canada: University of British Columbia, Department of Language Education.

Hyland, K. (2003). Genre-based pedagogies: Asocial response to process. *Journal of Second Language Writing, 12*, 17-29. https://doi.org/10.4236/ce.2015.66062

Hyland, K. (1990). A genre description of the argumentative essay. *RELC Journal, 21*, 66-78. https://doi.org/10.1177/003368829002100105

Hyland, K. (2007). Genre pedagogy: Language, literacy and L2 writing instruction. *Journal of Second Language Writing, 16*, 148-164. https://doi.org/10.1016/j.jslw.2007.07.005

Hudson, T. (1989). Mastery decisions in program evaluation. In R. K. Johnson (Ed.), *The second language curriculum* (pp. 259-269). Cambridge: Cambridge University Press. https://doi.org/10.4236/ojml.2015.55037

Lantolf, J. P., & Poehner, M. E. (2004). *Dynamic assessment in the language classroom*, (CALPER Professional Development Document CPDD-0411). University Park, PA: The Pennsylvania State University, Center for Advanced Language Proficiency Education and Research. https://doi.org/10.4236/ce.2022.133049

Lantolf, J. P., & Thorne, S. L. (2009). *Sociocultural theory and the Genesis of second language development* (pp.327-328) (4th ed.). Oxford: Oxford University Press. https://doi.org/10.4236/ce.2022.133049

Long, M. (1991). Focus on Form: A Design Feature in Language Teaching Methodology. In K. de Bot, R. Ginsberg, & C. Kramsch (Eds.), *Foreign Language Research in Cross-Cultural Perspective* (pp. 39-52). Amsterdam: Benjamins. https://doi.org/10.1075/sibil.2.07lon

Long, M., & G. Crookes. (1993). Units of Analysis in Syllabus Design: The Case for Task. In G. Crookes, & S. Gass (Eds.), *Tasks in Language Learning* (pp. 9-54). Clevedon: Multilingual Matters. https://doi.org/10.4236/msce.2017.57008

Moradian, M., Asadi, M., & Azadbakht, Z. (2019). Effects of concurrent group dynamic assessment on Iranian EFL learners' pragmatic competence: A case of requests and refusals. *Research in Applied Linguistics, 10*(2), 106-135. https://doi.org/10.4236/ojapps.2022.1210110

Poehner, M. E. (2005). *Dynamic assessment of oral proficiency among advanced L2 learners of French*. Unpublished doctoral dissertation, The Pennsylvania State University, University Park, PA. https://doi.org/10.4236/ce.2022.133049

Poehner, M. E. (2008). *Dynamic assessment: A Vygotskian approach to understanding and promoting L2 development*. United States: Springer. https://doi.org/10.4236/ce.2022.133049

Poehner, M. E. (2009). Group dynamic assessment: Mediation for the L2 classroom. *TESOL Quarterly, 43*(3), 471-491. https://doi.org/10.4236/ce.2022.133049

Poehner, M. E., & Lantolf, J. P. (2010). Vygotsky's teaching-assessment dialectic and L2 education: The case for dynamic assessment. *Mind, Culture, and Activity: An International Journal, 17*(4), 312-330. https://doi.org/10.4236/ce.2022.133049

Tajeddin, Z., & Tayebipour, F. (2012). The effect of dynamic assessment on EFL learners' acquisition of request and apology. *The Journal of Teaching Language Skills (JTLS), 4*(2), 88-118.

Torrance, H., & Pryor, J. (2002). *Investigating Formative Assessment: Teaching, Learning and Assessment in the Classroom*. Maidenhead: Open University Press. https://doi.org/10.4236/ojml.2019.91001

Torrance, H., & Pryor, J. (2001). Developing Formative Assessment in the Classroom: Using Action Research to Explore and Modify Theory. *British Educational Research Journal, 27*(5), 615-631. https://doi.org/10.1080/01411920120095780

Tzeng, G., & Huang, J. (2011). Multiple Attribute Decision Making Methods and Applications. CRC Press, Taylor and Francis Group, A Chapman & Hall Book, Boca Raton. Lantolf, J. P., (2009). *Sociocultural theory and second language learning* (pp.17-18) (5th ed.). Oxford: Oxford University Press. https://doi.org/10.4236/ce.2022.133049

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Massachusetts: Harvard University Press. https://doi.org/10.4236/ce.2021.127110

Vygotsky, L. S. (1986). *Thought and language*. Cambridge, MA: MIT Press. https://doi.org/10.4236/oalib.1101388

Appendix

Converting TOEFL Independent Argument Essay Writing Task Raw Score

Raw Score	Scaled Score
5.00	30
4.75	29
4.50	28
4.25	27
4.00	25
3.75	24
3.50	22
3.25	21
3.00	20
2.75	18
2.50	17
2.25	15
2.00	14
1.75	12
1.50	11
1.25	10
1.00	8
-	7
-	5
-	4
-	0

Copyrights