

Using Debate to Develop Writing Skills for IELTS Writing Task 2 among STEM Students

Daria Arzhadeeva, Natalia Kudinova

National Research University Higher School of Economics

Correspondence concerning this article should be addressed to Daria Arzhadeeva, National Research University Higher School of Economics, 20 Myasnitskaya Ulitsa, 101000, Moscow, Russia.
E-mail: darzhadeeva@hse.ru

The paper focuses on the issue of developing essay writing skills in the context of IELTS preparation and explores the issue of whether academic debate can enhance STEM students' ability to structure their essays, develop a smooth progression of ideas, and provide supported and extended arguments, which, in turn, may result in higher scores for the IELTS Task Response and Coherence and Cohesion categories. To answer this, a study was undertaken in the academic years 2016/2017 and 2017/2018 among STEM undergraduate students in the National Research University Higher School of Economics, Moscow, Russia. The study involved two groups of students (36 students in each): the group that attended regular IELTS preparation classes and the other that, in addition to regular classes, attended debate classes where among other things Toulmin's argument structure was taught. At the beginning and end of the experiment both groups submitted essays that were analysed according to IELTS rubrics for Task Response and Coherence and Cohesion, and the presence or absence of the elements of Toulmin's argument structure. In addition, the essays were assessed by an independent IELTS teacher. An independent-samples t-test and Levene's test were utilised to determine the significance of the collected data. The findings revealed that, on average, the students of the experimental group scored well in Task Response and Coherence and Cohesion, yet some results were inconsistent, which requires further research.

Keywords: debate, academic debate, STEM students, IELTS, Toulmin's structure, argumentation theory

Introduction

The International English Language Testing System (IELTS) is an English proficiency test that is taken by those who plan to work or study in an English-speaking environment. Comprising IELTS Academic and IELTS General Training, the test is accepted by more than 10,000 educational institutions worldwide, as the IELTS official website states.¹ Importantly, many universities in non-English speaking countries require an IELTS certificate as proof of one's command of the language (Lewthwaite, 2007; Sanonguthai, 2011).

IELTS Academic, which is the primary focus of this article, consists of four parts: Listening, Reading, Writing, and Speaking. Out of these four, a special challenge is posed by Writing. According to Test Taker Performance,² the mean band score of Academic Writing was the lowest compared to the other three parts across all 40 countries listed: 5.54 for males and 5.64 for females. The results in the Russia were not an exception, with the mean score being 5.97.

To meet international educational standards, the National Research University Higher School of Economics (HSE) in the Russian Federation designed the syllabus of its English course to accommodate the full format of IELTS Academic. As stated in *The Framework for Developing Foreign Languages Communicative Competence among Students of Bachelor's Programs and Specialist's Program* at the National Research University Higher

¹ IELTS. (2018a). IELTS for study. <https://www.ielts.org/what-is-ielts/ielts-for-study>

² IELTS. (2018b). Test taker performance 2018. <https://www.ielts.org/teaching-and-research/test-taker-performance>

School of Economics,³ by the end of the second year, students are to demonstrate English proficiency at B2-C1 CEFR level, that is, students have to gain IELTS Band 7 to get an excellent mark.

By year 2016, we had been preparing STEM (Science, Technology, Engineering, Mathematics) students for IELTS for two years and had noticed that our students tended to struggle with the Writing part of the exam, particularly with writing essays. We thought that such a situation could be explained by the fact that the language learning strategies employed by STEM students tend to differ from non-STEM students (Han, 2015). For instance, Cheng, Xu, and Ma (2007) found that STEM students use metacognitive strategies less frequently. The study by Han (2015) revealed that for writing skills STEM students “employ pre-writing strategies and attempt to find meanings of new words rather than construct messages in sentences before writing” (p. 96). On the other hand, according to González-Becerra (2019), the literature on this group is rather scarce, which means that more research into the peculiarities of language learning of STEM students is needed to tailor the courses for their needs.

The collected information motivated us to search for new ways to teach writing. Among them was debate structure (Sanongunthai, 2011). Therefore, we made the decision to carry out research to test the assumption that debate enhances the writing skills of STEM students. This paper presents the results of our study.

Theoretical Background

The Writing part of IELTS Academic comprises Task 1 and Task 2, with the Task 2 score being twice as important as that of Task 1.⁴ For our research, we focused primarily on Writing Task 2, which requires examinees to write an essay of around 250 words.⁵ As of August 18, 2018, on the IELTS official website it is stated that “[i]n Writing Task 2, test takers are given a topic to write about in an academic or semi-formal/neutral style. Answers should be a discursive consideration of the relevant issues”.⁶ What is of particular pertinence to our study is that “[t]his task assesses the ability to present a clear, relevant, well-organised argument, giving evidence or examples to support ideas and use language accurately”.⁷

Four criteria are used to assess essays: Task Response, Coherence and Cohesion, Lexical Resource, and Grammatical Range and Accuracy (see IELTS Scores Guide (2018) for the complete description of the rubrics). It is important to highlight that these four criteria contribute equally to the overall score of the writing part. In this study, we look in detail into Task Response (TR) and Coherence and Cohesion (CC) as it is these criteria, we hypothesize, that can be developed by means of debate.

TR refers to formulating an opinion about a given statement and supporting it with evidence and examples taken from one’s experience. To score well in TR, one should cover all of the points raised in the task question and provide well-developed and fully-supported arguments. The CC criterion involves the logical structuring of an essay and use of cohesive devices.⁸ To get the highest score in CC, “there needs to be a smooth progression of sentences and ideas” (Soodmand Afshar et al., 2017, p. 7). Interestingly, Coffin (2004) pointed out that “although argument structure is not explicitly part of the IELTS assessment criteria nor a focus for rater training, the overall band descriptors used in candidates’ reports make general reference to their ability to argue” (p. 233). This idea is later reiterated by Soodmand Afshar et al. (2017). Coffin (2004) goes on to recommend that English language students preparing for IELTS should be made aware of different argumentative structures.

Argumentative structures are typically studied within argumentation theory. In the framework of this theory there is still no consensus on the definition of the term ‘argumentation’ (Van Eemeren, 2015). For our purpose, we have chosen the definition suggested by Dupin de Saint-Cyr, Bisquert, Cayrol, and Lagasquie-Schiex:

Argumentation can be viewed as a process done in order to exchange information together with some justification with the aim to obtain well-justified knowledge, or with the aim to increase

³ The framework for developing foreign languages communicative competence among students of bachelor’s programs and specialist’s program in National Research University Higher School of Economics. (2017). National Research University Higher School of Economics.

⁴ British Council. (2018). IELTS Scores Guide.

⁵ IELTS. (2018c). Test format - Academic Writing. <https://www.ielts.org/about-the-test/test-format>

⁶ IELTS. (2018c). Test format - Academic Writing. <https://www.ielts.org/about-the-test/test-format>

⁷ IELTS. (2018c). Test format - Academic Writing. <https://www.ielts.org/about-the-test/test-format>

⁸ IELTS. (2018c). Test format - Academic Writing. <https://www.ielts.org/about-the-test/test-format>

or decrease the approval of a point of view, or to confront and combine different views. (Dupin de Saint-Cyr et al. 2016, p. 57)

In our view, this definition conveys the core of Task 2, i.e., to provide a point/points of view and support it/ them with evidence. Furthermore, such an understanding of argumentation is in accord with the structure and the objectives of debates.

Out of many possible argumentation structures (Coffin, 2004; Stab & Gurevych, 2014), the most practically useful and effective for both written argumentation (Connor & Mbaye, 2002) and debates (Arzhadeeva & Kudinova, 2017) is Toulmin's model of argument structure. According to Toulmin, Rieke, & Janik (1984), the structure of any argument consists of four main and two optional elements: claim, data/grounds, warrants, backing, rebuttal, and quantifier. Claim is understood as an assertion or "a well-defined position ... to consider and discuss" (Toulmin, Rieke, & Janik, 1984, p. 29). Data, or grounds, is construed as an underlying foundation that is required for a claim to be accepted as reliable. Normally, data includes statistical data, facts, the results of an experiment, etc. By warrant, Toulmin et al. (1984) mean "statements indicating how the facts on which we agree are connected to the claim or conclusion now being offered" (p. 45). However, warrants have to be validated and to that end backing is introduced. Backing is "the general body of information" (p. 26) utilized to make sure that warrants are reliable. The first optional element of the argument structure is rebuttal, which is viewed by Toulmin et al. (1984) as "the extraordinary or exceptional circumstances that might undermine the force of the supporting arguments" (p. 95). The other optional element is qualifier, which is a phrase or a word that shows the degree of certainty for the arguments used in a debate.

To illustrate, in an opinion essay the aforementioned elements of the argument structure work in the following way. Claim is the position of the writer usually expressed in an introduction. Data is a set of supporting arguments presented in the main body of an essay that include facts, examples, statistics, or expert opinions that students might know about. Warrant is normally a sentence that shows a logical connection between the main idea and each supporting argument. Backing is a fact that is used to confirm the validity of logical connections expressed in warrant. Rebuttal is a counterargument to the position of the writer frequently presented in a separate paragraph. Qualifier is a modal verb, adverb of certainty, or conditional sentence showing how certain the writer is about their position.

In their recent research, Ananda, Arsyad, and Dharmayana (2018) analysed sample essay answers with band score 8-9 from the publicly available IELTS materials from 2017⁹ and proved that all these elements are used to write an argumentative essay in IELTS Writing. They identified the most commonly utilized elements of argument in Task 2: claim, data, and warrant; and went on to point out the existence of two structures in the IELTS essay: the simple structure involving claim and data and the strong structure including "claim, data as obligatory elements, warrant as conventional element, backing, qualifier and rebuttal as the optional element" (Ananda et al., 2018, p. 11). Thus, to successfully write an IELTS essay in Task 2, one seems to need to use the strong argument structure.

We looked into a number of existing teaching methods that are believed to improve students' writing performance in order to find ones that help them develop the argument structure. Ostovar-Namaghi and Safaee (2017) conducted a data-driven study, providing a list of methods that, in teachers' opinions, are of particular help to students, among which are exposing candidates to sample answers, teaching grammar and vocabulary as a prerequisite to writing, teaching prefabricated patterns, raising candidates' awareness of scoring criteria, teaching discourse markers, and encouraging learners to develop their content knowledge. It is important to note that this list is based solely on the teachers' experiences. However, there is a significant body of studies conducted mostly by Iranian researchers who "have tested the effectiveness of theory-driven interventions under controlled experimental conditions to come up with universally applicable generalizations" (Ostovar-Namaghi & Safaee, 2017, p. 74). Among those studies are ones devoted to students' exposure to model essays (Abe, 2008; Bahgeri & Zare, 2009; Dickinson, 2013); teaching meta-discourse markers (Allami & Serajford, 2012); collaborative writing (Shehadeh, 2011); computer-assisted writing (Algraini, 2014; Bani Abdelrahman, 2013); using games (Ocriciano, 2016); corrective feedback (Ganji, 2009; Ketabi & Torabi, 2012; Sanavi & Nemati, 2014); using corpora or corpus-based materials (Smirnova, 2017); and debates (Sanongunthai, 2011).

⁹ About IELTS. Academic IELTS essays sample (2017). Retrieved from <https://ieltsmaterial.com/writing/>

USING DEBATE TO DEVELOP WRITING SKILLS FOR IELTS

From the intervention techniques listed above, we decided to focus only on debates, as during debates learners are supposed to utilize Toulmin's argument structure, which is also essential to writing a successful IELTS Writing Task 2.

In our study, we construe academic debate as

a teaching method which is aimed at developing and enhancing a learner's aspects of personality and performance such as communicative skills, problem solving skills, active listening, critical thinking, creativity, motivation, and adaptability and helps to gain knowledge and overcome stereotypes. (Kudinova & Arzhadeeva, 2020, pp. 4–5)

The prominent feature of academic debate is a structured procedure that is agreed upon beforehand: two opposing teams discuss a controversial topic by presenting arguments in favour and against it; and their performance is assessed by a group of judges.

There are a number of studies showing that debates have a positive impact on the following aspects: students' motivation (Schroeder & Ebert, 1983); communication skills (Alasmari & Ahmed, 2013; Temjakova & Ustinova, 2012); active listening, critical thinking, and creativity (Frankovskaya, 2002; Temjakova & Ustinova, 2012); adaptability (Gaumer Erickson & Noonan, 2016; Kudinova & Arzhadeeva, 2020). In addition, debates enable students to acquire knowledge (Khopiyajnen, 2015; Vo & Morris, 2006) and challenge stereotypes (Schroeder & Ebert, 1983). However, we have not found enough research on the influence of debate on students' performance during the IELTS exam. The only scholar who directly looked into the matter is Sanonguthai (2011), who explored the effectiveness of debate in English as a teaching method used to prepare high school students for the IELTS writing module, specifically Task 2. To that end, Sanonguthai chose 20 senior students who attended an academic writing course that was taught by a specially invited debate coach. The coach tailored the syllabus to the needs of the students. She combined debate sessions with activities aimed at extending their vocabulary and improving their grammar. Sanonguthai observed one class on week 8 and conducted follow-up interviews with both the students and the coach. One month later, the students took a mock IELTS test run by an experienced IELTS teacher. Three students' essays were randomly selected and studied, and the writers of those essays were interviewed. At the end of the study, Sanonguthai drew the following conclusions:

- debate is one of the most effective methods for teaching students to write and think more critically;
- debate hones students' skills to organize their written work and to structure an argument;
- debate extends students' background knowledge of the most common issues discussed in IELTS

However, after reading the research, certain points remained unclear. Sanonguthai's analysis does not provide information on the participants' backgrounds (initial level of English, gender, previous English learning experience, the length of IELTS preparation, previous debate experience), which could be significant as they might have affected the results. The description of the procedure also leaves room for speculation. Sanonguthai does not specify how long the debate course was and how many debate classes per week the students had. The degree of the teacher's and the students' involvement in class debate activities was unequal, as the coach appears to have been more actively engaged than her students. She provided them with content, rephrased their sentences, and restructured their arguments, whereas the students seem to have assumed a passive role. Only six out of 20 gave speeches and it is not entirely clear what the rest did during the debate itself. The conclusion drawn by the scholar from the fact that a few students failed to submit one essay and that this affected their final scores in writing appears to be far-fetched. In light of all this, a new more thorough study is needed to prove or disprove the conclusions Sanonguthai made about the effect of debates on students' performance in IELTS Writing Task 2. This is what the current study intends to do.

Research Hypotheses

In our study we tested the following two hypotheses:

1. Debate improves STEM students' ability to provide well-developed and fully-supported arguments, which can help them to score well in Task Response.
2. Debate enables STEM students to structure their essays well and develop a smooth progression of ideas, which can help them to score well in Coherence and Cohesion.

Methodology

Research Design

To test our hypotheses, in 2016 we organized a debate club called the ‘Debating Society’ in HSE University in Moscow, Russia. The club functioned from September to March because the IELTS -type exam was scheduled at the end of March. In the 2016/2017 academic year the participants attended weekly club meetings. In the 2017/2018 academic year, we decided to continue the work of the club, but with new participants. That was done to increase the sample size.

Settings

HSE University is the leading university in Russia.¹⁰ It offers IELTS preparation that starts in the first year, but it is quite sporadic and mostly limited to the IELTS Listening and Reading parts. Writing Task 2 comes into focus only in the second-year syllabus, which means that all students have one year of General English and start to intensively prepare for IELTS only in their second year.

Second-year students have a 90-minute English class twice a week; all in all, they have 52 classes from September to March. The classes are held in English and organised around *Complete IELTS Bands 6.5–7.5 Student’s Book* by Brook-Hart and Jakeman (2013).¹¹ In class, students are familiarised with the IELTS exam format, requirements, and assessment criteria. Apart from reading, listening, and speaking, the students are trained to do IELTS Writing Tasks 1 and 2. The classes are taught by CELTA-trained Russian speaking teachers with 9–12 years of teaching experience.

Preparation for IELTS Writing Task 2 involves discussing different essay structures, examining sample answers, and learning cohesive devices, subject-specific vocabulary, and grammar. Every second or third week, students write an essay following the required IELTS format.

Participants

At the beginning of the 2016/2017 academic year, 138 second-year HSE University undergraduates majoring in STEM were invited to take part in the debate club. The number of students who volunteered to come to the first club meeting was 19. After having the nature and purpose of the study explained to them, only 11 students agreed to participate and regularly attended the debate meetings (at least 90% of the classes) throughout the academic year: five male and six female students. In 2017/2018, out of 152 second-year students, 42 students agreed to attend debate classes. Yet, only 25 (15 males and 10 females) of them decided to take part in the study, whereas the rest attended debate classes but did not participate in the study. We decided to view the debate participants from both academic years (36) as an experimental group due to their similarity: they were all Russian-speaking second-year STEM students who followed the same curriculum, needed to meet the same English course requirements, and planned to take the IELTS exam in March. In the experimental group, the English level of the students, which is traditionally checked at the end of the first year when all first-year students sit for the HSE independent English exam,¹² was Intermediate - Upper-Intermediate, corresponding to B1–B2 CEFR (The Common European Framework of Reference for Languages). However, there were a few students whose level of English was Pre-Intermediate (A2). Importantly, all of the participants had learnt English since they were in grade 2 of primary school. Upon completion of the school English program, only one student took a standardised test (the Russian Unified State Exam; level B2 according to the Specification of Test Measures for the Unified State Exam in Foreign Languages¹³) and scored 96 points out of 100. None of the students took any of the internationally recognized tests (B2 First, C1 Advanced, C2 Proficiency, the TOEFL® test, IELTS, etc.). As for the length of IELTS preparation in HSE, it starts in the first year, but is quite sporadic

¹⁰ HSE Among Top Three Russian Universities. (2020, July 6). <https://www.hse.ru/en/news/edu/376039157.html>

¹¹ Brook-Hart, G., & Jakeman, V. (2013). *Complete IELTS Bands 6.5–7.5: Student’s Book*. Cambridge: Cambridge University Press.

¹² *The framework for developing foreign languages communicative competence among students of bachelor’s programs and specialist’s program in National Research University Higher School of Economics*. (2017). National Research University Higher School of Economics.

¹³ *Specification of test measures for the Unified State Exam in foreign languages*. (2018). Federal Institute for Educational Measurement (FIPI).

and mostly limited to the Listening and Reading parts of the IELTS exam. Writing Task 2 comes into focus only in the second-year syllabus, which means that all of the participants of the study had one year of General English and started to intensively prepare for IELTS only in their second year.

The participants of the control group were also the second-year Russian-speaking STEM students who attended regular English classes, but did not take part in debates. After having learnt about the purpose of the study, they volunteered to participate in it. Thus, the control group consisted of 36 students, 11 students (7 males and 4 females) from the 2016/2017 academic year and 25 students (12 males and 13 females) from the next year. The English level of the students in the control group was similar to that of the experimental group in addition to their language and IELTS preparation background.

Furthermore, we would like to underscore that out of the 36 students in the experimental group only one had a previous experience with Parliamentary debate. Those debates were held in the Russian language.

Procedure

In addition to the regular classes described in the Settings section, the students in the experimental group attended debate classes held once a week from September to March. Importantly, the teachers of the regular classes did not teach the debate classes. The latter were taught by two teachers with three years of debate teaching experience and ten years of English teaching experience.

A typical debate class entailed:

- theoretical input: the students were introduced to the rules of debates, Toulmin's structure of an argument, strong and weak arguments, propositions, rebuttal arguments, points of interest, typical fallacies, hedging, etc.;
- exercises to put theory into practice;
- mini-debates in pairs or groups of three/four (15-20 minutes long).

Every third class was a full-class debate. The topics of the debates covered issues that were common for the IELTS exam. They were normally announced one week before the debate.

Parliamentary debate was chosen as the debate format. In brief, this format involves two teams with opposite opinions debating a topic and each speaker delivering their speech within a certain time period. However, a few alterations were made to tailor the format to the needs and English level of the participants. In every debate, all of the participants were randomly assigned to three teams: Proposition, Opposition, and a team of judges. The Proposition team introduced the motion and spoke for it, providing relevant arguments, statistics, and examples. The Opposition team spoke against the motion countering the arguments of the Proposition and providing their own. The team of judges assessed the performance of the Proposition and Opposition teams and after some discussion choose the winner. Before debating, the participants had 20 minutes to develop a strategy and choose suitable arguments. During the debate, each participant in each group was to speak publicly for 3 or 4 minutes presenting their argument and/or counterargument. While someone was speaking, a member of the opposing team was allowed to ask questions (points of information). In addition, at the end of the debate each team had the chance to ask two or three questions to weaken their opponent's position. It is also important to underscore that the teacher's role during the debate class was limited to organizing, monitoring, and feedback without active involvement in the preparation, performance, or assessment.

At the beginning and the end of the course (in September and March), during regular lessons the students from both the experimental and control groups wrote an IELTS Task 2 essay as a pre- and post-test. The essay topic for both first and final classes was as follows:

Some people claim that not enough of the waste from homes is recycled. They say that the only way to increase recycling is for governments to make it a legal requirement. To what extent do you think laws are needed to make people recycle more of their waste?¹⁴

¹⁴ Cambridge IELTS 11 Academic. Student's book with answers: Authentic examination papers. (2016). Cambridge: Cambridge University Press.

For the purposes of the experiment, the topic was repeated in both academic years: 2016/2017 and 2017/2018.

The essays were collected and analysed according to the criteria, which will be discussed further. In addition, the final essays were graded by one independent teacher with 13 years of IELTS teaching experience. She used the IELTS Task 2 band descriptors for Task Response and Coherence and Cohesion. The Intraclass Correlation Coefficient for Task Response was .947 and for Coherence and Cohesion was .938. The decision to invite an independent expert was made due to the fact that IELTS examinees are officially given the overall score for both Writing Task 1 and Task 2 without specifying the particular score for an essay or scores for individual criteria, which makes it impossible to understand students' performance in Task Response and Coherence and Cohesion in Writing Task 2.

In the analysis of the essays, we relied on the fact that according to Transfer of International English Language Test Results to HSE¹⁵ an 'excellent' mark (8 out of 10) equals IELTS band 7, so in our study we utilized the writing assessment criteria for IELTS band 7 (only Task Response and Coherence and Cohesion). The additional criterion applied to assess the essays was the use of Toulmin's model (Table 1). To simplify the assessment process, we resorted to the binary system of grading where zero (0) stood for the absence of an element and/or its improper use and one (1) stood for the proper use of an element.

Table 1
Assessment Criteria

Task Response (TR) IELTS (Band 7)	Coherence and cohesion (CC) IELTS (Band 7)	Additional Criteria: Toulmin's Model of Argument Structure (TA)
TR1: all parts of the task are addressed;	CC1: information and ideas are organised logically;	TA1: claim;
TR2: a clear position is formulated;	CC2: clear progression throughout;	TA2: data;
TR3: main ideas are presented, extended, and supported;	CC3: a clear central topic is present within each paragraph;	TA3: warrant;
TR4: ideas are supported by relevant evidence;	CC4: a range of cohesive devices is used appropriately.	TA4: backing;
TR5: ideas are supported by relevant examples.		TA5: qualifier; TA6: rebuttal.

To evaluate the results, we acquired the gain score for the experimental group and the gain score for the control group and did an independent-samples t-test ($p < .05$) between the two sets of gain scores. We also used Levene's test to assess the homogeneity of variances drawn from different samples (Tables 3, 5, and 7). All calculations were made using IBM® SPSS® Statistics Subscription Build 1.0.0.1347 64-bit edition.

Results

The results of the experiment are presented in Table 2 through Table 7 and grouped according to the hypotheses. Interestingly, as can be seen from the tables showing descriptive statistics (Tables 2, 4, and 6), in many cases the standard deviation is larger than the mean, which can be attributed to the fact that we used the binary system of grading with the two elements - zero and one.

Hypothesis 1: Debate improves STEM students' ability to provide well-developed and fully-supported arguments, which can help them to score well in Task Response.

The results of the students' performance in Task Response are presented in Tables 2 and 3. Table 2 shows the mean values for pre- and post-tests of both experimental and control groups, the mean of the gain score for each criterion in Task Response and the mean for the expert's opinion. In Table 3, the results of Levene's test and an independent-samples t-test for each TR criterion separately are given. Importantly, in all but one case the results of Levene's test exceeded 0.05. This is why, to evaluate the results of TR1, TR2, TR3, and TR5, we used the results of the t-test with equal variances assumed; however, for TR4 we utilised the one with equal variances not assumed.

¹⁵ Transfer of international English language test results to HSE. (2018, September 7). <https://www.hse.ru/en/studyspravka/perezach/>

USING DEBATE TO DEVELOP WRITING SKILLS FOR IELTS

Table 2

Task Response: Mean Values of Pre- and Post-tests for Experimental and Control Groups, Mean Gain Score between Two Sets, and Mean Gain Score for Expert's Opinion

Variable	Groups ^a	M (SD) Pre-Test	M (SD) Post-Test	M (SD) Gain Score
TR1	experimental group	0.417 (0.5)	0.917 (0.28)	0.5 (0.609)
	control group	0.583 (0.5)	0.833 (0.378)	0.25 (0.604)
TR2	experimental group	0.75 (0.439)	1 (0.000)	0.25 (0.439)
	control group	0.833 (0.378)	0.917 (0.28)	0.08 (0.439)
TR3	experimental group	0.306 (0.467)	0.611 (0.494)	0.31 (0.525)
	control group	0.361 (0.487)	0.222 (0.422)	-0.14 (0.539)
TR4	experimental group	0.028 (0.167)	0.694 (0.467)	0.67 (0.478)
	control group	0.000 (0.000)	0.111 (0.319)	0.11 (0.319)
TR5	experimental group	0.167 (0.378)	0.778 (0.422)	0.61 (0.549)
	control group	0.194 (0.401)	0.278 (0.454)	0.08 (0.604)
TR Expert's Opinion	experimental group			6.61 (0.903)
	control group			5.31 (1.064)

Note. ^a n = 36

Table 3

Task Response: Results of Levene's Test and Independent-Samples T-Test

Variable		Levene's Test for Equality of Variances		T-Test for Equality of Means			95% CI	Cohen's d
		F	Sig.	t	df	p		
TR1	Equal variances assumed	0.69	0.409	1.749	70	0.085	[-0.035, 0.535]	0.412
TR2	Equal variances assumed	2.991	0.088	1.61	70	0.112	[-0.04, 0.373]	0.387
TR3	Equal variances assumed	0.175	0.677	3.368	70	0.001	[0.181, 0.708]	0.804
TR4	Equal variances assumed	25.271	0.000	5.801	70	0.000	[0.365, 0.747]	
	Equal variances not assumed			5.801	60.979	0.000	[0.364, 0.747]	1.378
TR5	Equal variances assumed	1.196	0.278	3.881	70	0.000	[0.257, 0.799]	0.918
TR Expert's Opinion	Equal variances assumed	1.341	0.251	5.612	70	0.000	[0.842, 1.77]	1.317

Note. ^a p < .05

A closer look at Table 3 reveals that the results for TR1 and TR2 do not demonstrate a statistically significant increase, while the results of the experimental group for TR3, TR4, and TR5 show a statistical difference when compared to the control group. Such results are supported by a 95% confidence interval (CI), which is greater than zero for all three variables, and the effect size (Cohen's *d*), which is large.

Table 3 also provides information about the results of the t-test for expert's opinion on students' performance in TR. The independent expert assessed the experimental group statistically higher than the control group, which is confirmed by the values of CI and Cohen's *d*.

Hypothesis 2: Debate enables STEM students to structure their essays well and develop a smooth progression of ideas, which can help them to score well in Coherence and Cohesion.

The results of the students' performance in Coherence and Cohesion are presented in Tables 4 and 5. Table 4 shows the mean values for the pre- and post-tests in both the experimental and control groups, the mean of the gain score for each criterion in CC, and the mean for the expert's opinion. Table 5 outlines the results of Levene's test and an independent-samples t-test for each CC criterion separately, including a 95% CI and the effect size. Similar to the results of Levene's test for TR criteria, there is one case where the results of the test did not exceed 0.05 - CC4. Consequently, for this criterion we used the results of the t-test with equal variances not assumed.

Table 4

Coherence and Cohesion: Mean Values of Pre- and Post-tests for Experimental and Control Groups, Mean Gain Score between Two Sets, and Mean Gain Score for Expert's Opinion

Variable	Groups ^a	M (SD) Pre-Test	M (SD) Post-Test	M (SD) Gain Score
CC1	experimental group	0.25 (0.439)	0.639 (0.487)	0.39 (0.645)
	control group	0.278 (0.454)	0.361 (0.487)	0.08 (0.649)
CC2	experimental group	0.167 (0.378)	0.389 (0.494)	0.22 (0.591)
	control group	0.111 (0.319)	0.278 (0.454)	0.17 (0.561)
CC3	experimental group	0.222 (0.422)	0.722 (0.454)	0.5 (0.561)
	control group	0.472 (0.506)	0.611 (0.494)	0.14 (0.639)
CC4	experimental group	0.194 (0.401)	0.806 (0.401)	0.61 (0.494)
	control group	0.25 (0.439)	0.667 (0.478)	0.42 (0.692)
CC Expert's Opinion	experimental group			6.11 (1.141)
	control group			5.17 (1)

Note. ^a n = 36

Table 5

Coherence and Cohesion: Results of Levene's Test and Independent-Samples T-Test

Variable		Levene's Test for Equality of Variances		T-Test for Equality of Means			95% CI	Cohen's d
		F	Sig.	t	df	p		
CC1	Equal variances assumed	1.825	0.181	2.004	70	0.049	[0.001, 0.61]	0.479
CC2	Equal variances assumed	0.49	0.486	0.409	70	0.684	[-0.215, 0.326]	0.087
CC3	Equal variances assumed	0.437	0.511	2.548	70	0.013	[0.078, 0.644]	0.599
CC4	Equal variances assumed	7.052	0.01	1.372	70	0.174	[-0.088, 0.477]	
	Equal variances not assumed			1.372	63.356	0.175	[-0.089, 0.478]	0.316
CC Expert's Opinion	Equal variances assumed	0.251	0.618	3.735	70	0.000	[0.44, 1.449]	0.876

Note. ^a p < .05

As can be seen from Table 5, the difference between the groups was significant in two (CC1 and CC3) out of the four criteria. The 95% confidence intervals of [0.001, 0.61] for CC1 and [0.078, 0.644] for CC3 confirm that. Regarding the effect size, Cohen's *d* for CC1 was small, below 0.50, and Cohen's *d* for CC3 was medium, greater than 0.50 but lower than 0.80.

Looking at the independent expert's score, we can see that the mean value of the control group is statistically lower than that of the experimental group, a 95% CI is greater than zero, and the effect size is large.

Additional Criteria: Toulmin's Model of Argument Structure (TA)

Tables 6 and 7 provide information on the students' performance in each structural element of Toulmin's model. Table 6 shows the mean values for the pre- and post-tests in the experimental and control groups and the gain score for each element of Toulmin's structure. Table 7 presents the results of the Levene's test and the t-test for each element. For TA4, the results of the Levene's test did not exceed 0.05, so we interpreted the data based on the results of the t-test with equal variances not assumed.

Table 6

Toulmin's Model of Argument: Mean Values of Pre- and Post-tests for Experimental and Control Groups, and Mean Gain Score between Two Sets

Variable	Groups ^a	M (SD) Pre-Test	M (SD) Post-Test	M (SD) Gain Score
TA1	experimental group	0.722 (0.454)	0.861 (0.351)	0.14 (0.543)
	control group	0.833 (0.378)	0.75 (0.439)	-0.08 (0.649)
TA2	experimental group	0.444 (0.504)	0.778 (0.422)	0.33 (0.632)
	control group	0.694 (0.467)	0.278 (0.454)	-0.42 (0.649)
TA3	experimental group	0.083 (0.28)	0.667 (0.478)	0.58 (0.5)
	control group	0.167 (0.378)	0.389 (0.494)	0.22 (0.637)
TA4	experimental group	0.000 (0.000)	0.556 (0.504)	0.56 (0.504)
	control group	0.000 (0.000)	0.167 (0.378)	0.17 (0.378)
TA5	experimental group	0.222 (0.422)	0.75 (0.439)	0.53 (0.736)
	control group	0.222 (0.422)	0.528 (0.506)	0.31 (0.668)
TA6	experimental group	0.222 (0.422)	0.056 (0.232)	-0.17 (0.507)
	control group	0.222 (0.422)	0.167 (0.378)	-0.06 (0.583)

Note. ^a n = 36

TA1: claim; TA2: data; TA3: warrant; TA4: backing; TA5: qualifier; TA6: rebuttal.

Table 7

Toulmin's Model of Argument: Results of Levene's Test and Independent-Samples T-Test

Variable		Levene's Test for Equality of Variances		T-Test for Equality of Means			95% CI	Cohen's d
		F	Sig.	t	df	p		
TA1	Equal variances assumed	0.59	0.445	1.576	70	0.12	[-0.059, 0.503]	0.368
TA2	Equal variances assumed	0.18	0.672	4.965	70	0.000	[0.449, 1.051]	1.171
TA3	Equal variances assumed	0.276	0.601	2.674	70	0.009	[0.092, 0.63]	0.629
TA4	Equal variances assumed	25.221	0.000	3.704	70	0.000	[0.179, 0.598]	
	Equal variances not assumed			3.704	64.911	0.000	[0.179, 0.599]	0.876
TA5	Equal variances assumed	0.395	0.532	1.341	70	0.184	[-0.108, 0.553]	0.313
TA6	Equal variances assumed	0.001	0.974	-0.863	70	0.391	[-0.368, 0.146]	0.201

Note. ^a p < .05

The significant difference between the experimental and control groups can be found in TA2, TA3, and TA4. The 95% CI for all three variables are greater than zero; the effect size for TA3 is medium while that for TA2 and TA4 is large.

Discussion

Hypothesis 1: Debate improves STEM students' ability to provide well-developed and fully-supported arguments, which can help them to score well in Task Response.

The findings show that among the students who attended the debate classes there was a far larger proportion of those who presented and extended main ideas in their essays and supported them with relevant evidence and examples (TR3, TR4, and TR5), and the difference between the groups was consistent and substantial (Table 3). Such an outcome may be explained by the fact that during debates the participants were taught to clearly formulate the core ideas and develop them by giving relevant statistics, facts, examples, or experts' opinions. These skills correspond to the elements of data (TA2) and warrant (TA3) within the framework of Toulmin's model of argument structure, which was taught during the debate classes. In the essays of the students who attended debates, data and warrant were used more frequently than in the essays of those who attended only the regular classes (Table 7).

However, participation in debates, seemingly, did not bring any changes to the students' ability to fully address all parts of the task (TR1) and to formulate a clear position in their essays (TR2) (Table 3), which is consistent with the results obtained for the elements of Toulmin's structure. There was no significant difference in the use of claim (TA1) between the two groups (Table 7). The conducted tests did not allow us to say whether all participants of the study managed to perform well or poorly in TR1, TR2, and TA1. However, we might suggest that if they had all performed well, that could be attributed to the fact that these aspects are traditionally focused on in regular IELTS preparation classes.

Although debate had a considerable impact on only three out of five criteria in Task Response, we can still see a positive trend in the way debates enabled students to perform better on their essays. Such results are also confirmed by the independent expert's grading, which showed that the students who attended the debate classes performed better in Task Response than those who only attended regular IELTS preparation classes. This allows us to say that the results of our study support the hypothesis that participation in debate considerably enhances STEM students' ability to give well-developed and fully-supported arguments, helping them to score well in Task Response.

Hypothesis 2: Debate enables STEM students to structure their essays well and develop a smooth progression of ideas, which can help them to score well in Coherence and Cohesion.

Although the results of the t-test demonstrated that the students from the experimental group were better able to organise their thoughts logically (CC1), that difference was trivial (Cohen's $d = 0.479$). Such an outcome does not prove that debates helped students hone this skill. The other aspect that showed a significant improvement was CC3 (a clear central topic being within each paragraph) and, based on the medium effect size, the change was likely to be visible. Regarding CC2 (having a clear progression throughout the essay) and CC4 (using a range of cohesive devices appropriately), the debate classes did not give the students who attended them a competitive advantage over those who did not. Yet according to the independent expert, the students who attended the debate classes wrote more coherent essays and the difference between the groups was substantial.

Interpreting the results for the ability to use a range of cohesive devices appropriately, the lack of a significant difference between the control and the experimental groups can be attributed to two facts: first, cohesive devices are normally studied as a compulsory part of the IELTS preparation course at HSE University; second, the cohesive devices used in debates are more typical of spoken, not written, discourse. As to ensuring a clear progression throughout the essay, it is possible that the students of the experimental group did not differ from those of the control group due to the aspects of debate organisation. In debates, each student puts forward one logically clear and concise argument, not a string of arguments typical of essays. Constructing a clear progression within a range of arguments is a team task, as during the debates members of the team contribute equally to the defense of their main position. That aspect must have resulted in the students of the experimental group being better able to select one central topic within each paragraph. Regarding CC1, it is likely that regular debate attendees improved their ability to organise their thoughts logically; however, they might have failed to apply the acquired skill to the new circumstances.

Taking into consideration the results of the independent expert's assessment, it might be said that the overall findings seem rather inconclusive and contradictory. This might be accounted for by the fact that participation in debates did not have a considerable impact on each separate criterion within the Coherence and Cohesion criteria of IELTS Writing Task 2; however, it affected the overall impression of an essay's logical structure. We might also suggest that what contributed to the better structured essays in the experimental group was the presence of such elements as warrant (TA3) and backing (TA4) of Toulmin's strong structure (Table 7).

Based on the findings in Coherence and Cohesion, we cannot definitely state that participation in debates improves STEM students' ability to structure their essays well and develop a smooth progression of ideas that can help them to receive a higher score in Coherence and Cohesion. It follows that more research is needed into the influence of debate on the Coherence and Cohesion criteria of IELTS Writing Task 2.

Limitations of the Present Study

We would like to point out a number of limitations to the current study. The first one is the rather small sample. The small size of the sample is explained by the fact that participation in the debates as well as in the study was

voluntary. We attempted to increase it by considering the results of the two academic years as one sample. The second limitation is the lack of randomisation during the assigning of participants to the control and experimental groups, which means that we cannot be sure that the students who volunteered to attend debate classes did not have any unknown significant differences from the students of the control group. For example, the students of the experimental group might have been more interested in argumentation theory or have had additional English classes outside the university. Although this limitation is significant and might have affected the results, our sample was rather homogenous in other aspects: age, level of English, educational background, IELTS preparation experience, native language, etc. Furthermore, this limitation was difficult to overcome due to the extracurricular nature of debate classes, students' schedules, and university rules.

Conclusions, Implications of the Study, and Directions for Future Research

In the current paper, we focused on preparing for IELTS Writing Task 2 through debate. We argued that the ability to apply Toulmin's model of argument structure may help STEM students score well in Task Response and Coherence and Cohesion. As a result of the study, the hypothesis concerning the impact of debates on Task Response was provisionally supported whereas the hypothesis regarding Coherence and Cohesion leaves room for speculation, even though the overall score for these criteria has improved.

The findings of the current study may be used in the field of ESL as one of the methods of test preparation, particularly for STEM students, for IELTS Writing Task 2. The use of debates may give STEM students deeper insight into how to structure and develop their ideas and provide extended arguments in order to score more highly in Task Response and Coherence and Cohesion. Apart from that, debate may introduce variety to the exam preparation classes and break the routines.

However, since there are a number of significant limitations to the current study, further research is needed to establish a clearer and more conclusive correlation between the use of Toulmin's model and IELTS scores for Task Response and Coherence and Cohesion. To refine our findings, we believe it would be necessary to carry out another study randomising the sample and expanding its size. Overall, these further studies might ascertain how replicable the currently collected results are and allow for more validity of the method.

Conflict of interests

The authors declare that they have no conflict of interest.

References

- Abdelrahman, O. N. M. B. (2013). The Impact of Using the Word Processor to Develop EFL Learners' Writing Skill at Al-Imam Mohammad Ibin Saud Islamic University. *IUG Journal of Humanities Research*, 21. 1–26. <https://doi.org/10.12816/0013730>
- Abe, M. (2008). Exploring the Role of Model Essays in the IELTS Writing Test: A Feedback Tool. (Master's thesis). The University of Queensland, Australia.
- Alasmari, A., & Ahmed, S. S. (2013). Using debate in EFL classes. *English Language Teaching*, 6(1), 147–151. <https://doi.org/10.5539/elt.v6n1p147>
- Algraini, F. (2014). The Effect of Using Padlet on Enhancing EFL Writing Performance. (Master's thesis). Al-Imam Muhammad Ibn Saud Islamic University, Kingdom of Saudi Arabia.
- Allami, H., & Serajfard, H. (2012). Engagement Markers: A Technique for Improving Writing Skills. *Journal of Language, Culture, and Translation*, 1, 71–83.
- Ananda, R. P., Arsyad, S., & Dharmayana, I. W. (2018). Argumentative features of International English Language Testing System (IELTS) essays: A rhetorical analysis on successful exam essays. *International Journal of Language Education*, 2(1), 1–13. <https://doi.org/10.26858/ijole.v2i1.4768>

- Arzhadeeva, D., & Kudinova, N. (2017). Razvitie kommunikativnih umenii kak odnoi iz sostavlyajushchih socialnoj kompetentnosti posredstvom tekhnologii 'debati' [Developing communication skills through debates]. *Philology at Moscow State Institute of International Relations (MGIMO) Journal*, 2(10), 55–64.
- Bahgeri, M., & Zare, M. (2009). The Role of Using IELTS Model Essays in Improving Learners' Writing and their Awareness of Writing Features. *Journal of English language Studies*, 1(1), 115–130.
- Cheng, X., Xu, K., & Ma, Y. (2007). A survey of engineering student's use of English language learning strategies. *The Journal of Asia TEFL*, 4(2), 123–140.
- Coffin, C. (2004). Arguing about how the world is or how the world should be: The role of argument in IELTS tests. *Journal of English for Academic Purposes*, 3, 229–246. <https://doi.org/10.1016/j.jeap.2003.11.002>
- Connor, U., & Mbaye, A. (2002). Discourse approaches to writing assessment. *Annual Review of Applied Linguistics*, 22, 263–278. <https://doi.org/10.1017/S0267190502000144>
- Dickinson, P. (2013). A Genre-based Approach to Preparing for IELTS and TOEFL Essay Writing Tasks. *Journal of Niigata University of International and Information Studies*, 16, 1–9.
- Dupin de Saint-Cyr, F., Bisquert, P., Cayrol, C., & Lagasquie-Schiex, M.-C. (2016). Argumentation update in YALLA (Yet Another Logic Language for Argumentation). *International Journal of Approximate Reasoning*, 75, 57–92. <https://doi.org/10.1016/j.ijar.2016.04.003>
- Frankovskaya, O. (2002). Razvitie refleksii u shkol'nikov cherez debatnuyu tekhnologiyu [Developing reflection among schoolchildren through debate]. https://experiment.lv/rus/biblio/debate/1_frankovskaja.htm
- Ganji, M. (2009). Teacher-correction, Peer-correction and Self-correction: Their Impacts on Iranian Students' IELTS Essay Writing Performance. *The journal of Asia TEFL*, 6(1), 117–139.
- Gaumer Erickson, A., & Noonan, P. (2016). Teacher guide college and career competency: adaptability. College and Career Competency Framework. <https://www.researchcollaboration.org/uploads/TeacherGuide-Adaptability.pdf>
- González-Becerra, I. (2019). Language learning among STEM students: Motivational profile and attitudes of undergraduates in a UK institution. *The Language Learning Journal*, 47(3), 385–401. <https://doi.org/10.1080/09571736.2017.1320421>
- Han, T. (2015). Estrategias de aprendizaje de Lenguas Extranjeras en el contexto de la educación STEM [Foreign language learning: Strategies in the context of STEM education]. *GIST Education and Learning Research Journal*, 11, 79–102. <https://doi.org/10.26817/16925777.287>
- Ketabi, S., & Torabi, R. (2012). Teaching academic writing in Iranian EFL classrooms: Teacher-initiated comments or peer-provided feedback? *Research in English Language Pedagogy*, 1(2), 58–65.
- Khojiyajnen, O. A. (2015). Tekhnologiya debatov v obuchenii inostrannym yazykam [Debate technology in teaching foreign languages]. *Vestnik YuGU*, 1(36), 158–160.
- Kudinova, N., & Arzhadeeva, D. (2020). Effect of debate on development of adaptability in EFL university classroom. *TESOL Journal*, 11(1), 1–13. <https://doi.org/10.1002/tesj.443>
- Lewthwaite, M. (2007). Teacher and student attitudes to IELTS Writing tasks: Positive or negative washback? *UGRU Journal*, 5, 1–23.
- Ocriciano, M. (2016). IELTS Writing: A gamification journey. *Cambridge English: Research Notes*, 64, 31–36.
- Ostovar-Namaghi, S.A., & Safaee, S.E. (2017). Exploring techniques of developing writing skill in IELTS preparatory courses: A data-driven study. *English Language Teaching*, 10(3), 74–81. <https://doi.org/10.5539/elt.v10n3p74>
- Sanavi, R., & Nemati, M. (2014). The Effect of Six Different Corrective Feedback Strategies on Iranian English Language Learners' IELTS Writing Task 2. *SAGE Open*, 4(2), 1–9. <https://doi.org/10.1177/2158244014538271>
- Sanonguthai, S. (2011). Teaching IELTS Writing module through English debate: A case study in Thailand. *Language Testing in Asia*, 1(4), 39–99. <https://doi.org/10.1186/2229-0443-1-4-39>
- Schroeder, H. & Ebert, D. G. (1983) Debates as a Business and Society Teaching Technique, *The Journal of Business Education*, 58 (7), 266–269. <https://doi.org/10.1080/00219444.1983.10534907>
- Shehadeh, A. (2011). Effects and student perceptions of collaborative writing in L2. *Journal of Second Language Writing*, 20(4), 286–305. <https://doi.org/10.1016/j.jslw.2011.05.010>
- Smirnova, E. (2017). Using Corpora in EFL Classrooms: The Case Study of IELTS Preparation. *RELJ Journal*, 48(3), 302–310. <https://doi.org/10.1177/0033688216684280>
- Soodmand Afshar, H., Movassagh, H., & Radi Arbabi, H. (2017). The interrelationship among critical thinking, writing an argumentative essay in an L2 and their subskills. *The Language Learning Journal*, 45(4), 1–13. <https://doi.org/10.1080/09571736.2017.1320420>
- Stab, C., & Gurevych, I. (2014). Identifying argumentative discourse structures in persuasive essays. In *Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)* (pp. 46–56). Publisher. <https://doi.org/10.3115/v1/D14-1006>

- Temjakova, T. V., & Ustinova, N. V. (2012). Tekhnologia “Parlamentsckie debaty” v vysshem obrazovanii [Parliamentary debate in tertiary education]. *Konzept*, (5), 48–52.
- Toulmin, C., Rieke, R., & Janik, A. (1984). *An introduction to reasoning*. Macmillan Publishers.
- Van Eemeren, F. H. (2015). Bingo! Promising developments in argumentation theory. In F. H. Van Eemeren & B. Garssen (Eds.), *Reflections on theoretical issues in argumentation theory* (pp. 3–25). Springer.
- Vo, H. X., & Morris, R. L. (2006). Debate as a tool in teaching economics: Rationale, technique, and some evidence. *Journal of Education for Business*, 81(6), 315–320. <https://doi.org/10.3200/JOEB.81.6.315-320>