

Article

Measuring the Effectiveness of Writing Center Consultations on L2 Writers' Essay Writing Skills

Shanthi Tiruchittampalam ^{1,*}, Alistair Ross ², Elizabeth Whitehouse ³ and Tom Nicholson ⁴

¹ Office of The Provost, United Arab Emirates University, POB 17172, Al Ain, UAE

² English Language Unit, University College, United Arab Emirates University, POB 17172, Al Ain, UAE; a.malcolm@uaeu.ac.ae

³ Student Academic Success Program, University College, United Arab Emirates University, POB 17172, Al Ain, UAE; ewhitehouse@uaeu.ac.ae

⁴ Institute of Education, College of Humanities and Social Sciences, Massey University, Private Bag 102904 North Shore, Auckland 0745, New Zealand; tricholson@massey.ac.nz

* Correspondence: shanthi.t@uaeu.ac.ae; Tel.: +9713-713-5963

Received: 5 August 2017; Accepted: 6 February 2018; Published: 19 February 2018

Abstract: With the international growth of English-medium education, tertiary institutions are increasingly providing academic support services to L2 students, and thus, the number of writing centers working with L2 student writers has also increased. Writing center practices originated in L1 English educational contexts and their appropriateness for L2 English writers requires examination. This study investigated the effect of writing center consultations on the essay writing skills of L1 Arabic foundation level students at an English-medium university in the Gulf region. Analysis was based on quantitative measures of writing ability of two distinct groups of students: an experimental group who participated in tutoring sessions at the university's writing center and a control group who did not. Findings indicated that students who participated in writing center consultations scored significantly higher in overall essay writing scores, as well as in two aspects of writing: task fulfilment (that is ideas) and text organization/coherence. These findings contribute to a limited bank of similar empirical studies on effectiveness of writing center sessions on students' essay writing ability. They also support the case for the expansion of writing center work beyond the domains of predominantly L1 English academic communities.

Keywords: writing center; student writing; L2 writers; essay writing skills; writing consultations

1. Introduction

Since writing centers became a common component of learning support services on United States (U.S.) college campuses, scholars have called for rigorous research to investigate the effectiveness of their practices (e.g., North 1984b; Driscoll and Perdue 2012). In recent years, the number of writing centers operating in English-medium institutions outside the U.S. has increased significantly (Rafoth 2014; Hyland 2018). Hence, there is a growing number of writing center practitioners who work almost exclusively with L2 English student writers. With this growth, the need for site-specific research into the appropriateness and effectiveness of writing center practices in non-traditional writing center contexts has come to the fore. The research presented here examines the efficacy of a writing center at an English-medium university in the Gulf region, focusing on the impact of one-to-one writing consultations on the essay writing skills of L1 Arabic students enrolled in the university's foundation program. The present report addresses first, the international growth of writing centers and the theoretical case for the efficacy of writing center consultations, followed by a review of previous studies into the efficacy of writing center consultations and a rationale for writing assessment. The study at hand is then detailed and its findings discussed.

2. Writing Centers: International Growth and Tutoring Methodology

Given impetus by open admission policies and broader student intake at U.S. higher education institutions in the 1960s, writing centers had become established as a common component of learning support services on U.S. college campuses by the 1990s. Since then there has been a significant growth in writing center work outside the U.S., a concomitant of expansion in education sectors globally and in particular, the increasing numbers of students enrolled in English-medium tertiary studies (Rafoth 2014). This growth led the U.S.-based National Writing Centers Association to rename itself the International Writing Centers Association in 1998 (Thaiss 2012). This body now oversees three affiliate regional writing center associations, including the Middle East and North Africa Writing Center Alliance, which serve predominantly non-L1 English regions of the world (International Writing Centers Association <http://writingcenters.org/affiliates>). Similarly, The Writing Center Directory (<http://web.stcloudstate.edu/writeplace/wcd/index.html>) records the presence of writing centers in 66 countries worldwide. These developments are associated with a growing body of students “from around the world (who) turn to writing centers to learn to write” (Rafoth 2014, p. 18). In the Gulf region, higher education institutions started establishing writing centers from the beginning of the 21st century, and since then they have proliferated (Eusafzai 2018). Hyland (2018, p. viii) comments that “the writing centre idea [. . .] has firmly taken hold among the universities of the Arabian Gulf”.

In developing the academic literacy and writing skills necessary for academic success, twenty-first century students have a wide array of online and technology-based resources available to them beyond classroom instruction (Rafoth 2014). However, the one-to-one tutorial service provided by writing centers offers students a unique and potentially optimal opportunity to improve their writing skills. The prime aim of writing centers is to help students become “better writers” (North 1984a, p. 438). Harris (1988) outlined the principles of writing center practice as follows. Firstly, writing center tutors provide student writers with individualized help in the form of one-to-one consultations. These consultations are student-led, with writers typically bringing a sample of their writing, establishing a goal for the session, and working collaboratively with the tutor to improve their work. The tutor–tutee relationship is non-hierarchical, thereby enabling students to raise concerns in a relaxed and non-judgmental environment and retain control over writer decisions. Tutors aim to help writers make long-term gains in writing skills by helping them improve their writing process (discussed below). They also play a key role in providing reader feedback, which helps students acculturate to the norms of their target discourse community.

The writing process is commonly understood to comprise prewriting, drafting, revising and editing stages. In her review of the psychology of writing development, Camp (2012) elaborates on these stages as being hierarchical, but non-linear and operating at different levels. For instance, at one level it involves cognitive maturity and refinement of ideas. At another level, Camp (2012) draws from socially-situated theories of writing development, which view writing as highly contextual and reflecting the shared values and conventions of the writer’s community, and the social interactions of the writer. Finally, the development of the writing process is steered by personal motivations of the individual writer. A more operational view of the writing process, outlined by Flower and Hayes (1981), involves the writer moving from abstract goals of focusing on topic knowledge and relevant literature to the generation of ideas in response to a writing prompt and then to the more concrete stages of writing and editing. Reigstad and McAndrew (1984, pp. 11–19) term the abstract goals higher order concerns (HOCs), commonly understood as relating to content, argument, textual organization and cohesion, in contrast to lower order concerns (LOCs), namely sentence level lexical and grammatical accuracy. Flower and Hayes (1981) also view writing as a recursive process, in which evaluation occurs, leading to sub-goals including forming clearer connections between ideas, reformulation and editing. In one-to-one writing consultations, tutors are ideally placed to encourage students to adopt these evaluative and recursive practices as part of the writing process.

Nordlof (2014) draws on Vygotsky’s Zone of Proximal Development, concept, as well as Puntambekar and Hübscher (2005) elaboration of scaffolding theory, to propose a theoretical framework

for understanding how writing center consultations facilitate the learning and development of writing skills. A Vygotskian approach views learning as taking place socially through interaction between a learner and a more capable second party, and in stages, whereby knowledge is first externally articulated and then gradually internalized by the learner. For Nordlof (2014), the one-to-one, dialogic and staged nature of a writing center consultation is highly conducive to this mode of learning. Closely connected is the concept of scaffolding, that is, techniques used by a second party to create conditions in which a learner can achieve their full learning potential at any one time. Puntambekar and Hübscher (2005, pp. 2–3) break down the concept of scaffolding into four components, namely “intersubjectivity”, “ongoing diagnosis”, “dialogic and interactive” instruction, and “fading”. Nordlof (2014) shows that each of these components can be mapped to specific stages of a writing consultation and the type of tutor–tutee interaction taking place at any particular time. For example, at the beginning of a consultation the tutor and tutee establish a shared purpose and collaborative working relationship, which constitutes intersubjectivity (Nordlof 2014, p. 56). In sum, the individualized, dialogic, and staged nature of a writing center consultation creates, potentially, an environment that is highly conducive to learning.

In addition to facilitating cognitive growth, writing center tutors also target motivational issues that can affect writers’ performance (Nordlof 2014, p. 57). Learner attributes known to correlate with writing ability, namely attitudes to writing and self-efficacy (a student’s beliefs about his or her capabilities as a writer), have been investigated for writing center effect (Babcock and Thonus 2012, pp. 160–61). For example, a semester long experimental study of students enrolled in eight first-year writing classes at a U.S. college found that writing center consultations had a positive effect on students’ attitudes to writing in that they became less apprehensive (Davis 1988). In another experimental study (Schmidt and Alexander 2012), the self-efficacy ratings of U.S. college students who participated in at least three writing center consultations during a semester improved markedly compared to those who did not. Thus, it seems that writing centers can bring about writer growth not only in terms of cognitive gains but also in attitudinal factors and perceptions of writing competence.

3. Studies on the Effects of Writing Center Consultations on Students’ Writing Ability

As the purpose of this study was to evaluate the effects of writing center tutorials on the writing ability of students, a review of the existing literature was conducted to locate studies similar to this one with quantitative methods that examine the relationship between writing center tutorials and writing ability. Of the ten quantitative studies found, all were conducted at tertiary institutions in the U.S. and the majority found positive results for writing center efficacy. The relevant findings of these studies are summarized below. However, none of the studies examined the link between writing center tutorials and writing abilities in L1 Arabic contexts. Furthermore, none used random assignment of students to experimental and control groups to control for the confounding factor of participant variability. In contrast, random assignment was chosen as part of the design for the present study.

Seven studies examined the impact of writing center tutorials on freshmen students’ overall course grades. Sutton and Arnold (1974) set up experimental and control groups of university freshmen at basic composition level. The experimental group attended one-to-one tutorials at the writing center, while the control group attended freshmen composition classes. Comparison of the participants’ grade point averages revealed that the experimental group received significantly higher grades at graduation. Similarly, Sandlon (1980) showed that students who received only writing center tutorials subsequently gained similar or higher grades than their freshman composition class counterparts. In contrast, Roberts (1988) found a non-significant effect for writing center courses as a replacement for classroom-based courses. Inconclusive results were also reported by Lerner (1998) when comparing the composition course grades of freshmen who attended writing center tutorials with the grades of those who did not. However, the basis of this study is subject to challenge, as Lerner used the students’ standardized college entry verbal test scores to establish the equivalence of the control

and intervention groups, and students self-selected their enrolment in the intervention group by choosing to visit the writing center. Conversely, Niiler (2003, 2005) quantitative studies did show a significant improvement in the grades of students who received writing center tuition on drafts of writing assignments. However, in the same manner as Lerner (1998), students self-selected for the experimental group by choosing to receive writing center tuition, which could have affected the results of both these studies. This is because the experimental group may then have contained students with higher motivation levels than the control group. To control this variable of students self-selecting writing center tuition, Irvin (2014) asked professors to require such tuition for 15 out of 28 classes from a range of courses. He established a “baseline” of each professor’s historical average of C grades or better and then compared the grades of study participants with these baselines. These results were positive for writing center tutorial effect, though again this method of measurement does not address the variables of group equivalence or rater variability. Thus, while most of the studies reviewed so far found a significant positive effect for writing center tuition in terms of student course grades, the impact of the tuition on specific aspects of student writing ability was not investigated. A further point of difference among these studies was that some examined the effect of writing center tuition as a replacement for, rather than supplementary to, composition classes.

Two other studies with freshmen as participants focused solely on the impact of writing centers on students’ grammatical awareness. Wills (1984) reported a significant increase in freshman ability to identify grammatical errors after writing center tutorials. Similarly, David and Bubloz (1985) measured the impact of writing center sessions on the writing of freshmen who had failed in composition and found significant improvements in grammatical accuracy. The findings of these studies were encouraging, but an assessment of writing ability cannot be limited to this one aspect of grammatical awareness.

The remaining study by Henson and Stephenson (2009) involved students from both basic and advanced composition courses, from freshmen level to senior level, with an experimental group who chose to attend additional writing center tutorials and a control group who did not, again allowing participant self-selection in the experimental group. Students from both groups wrote pre-test and post-test essays which were then assessed by the researchers, using a rubric to quantify different ratings for HOCs and LOCs of writing. Results indicated the experimental group had achieved significant increases in ratings for the HOCs, such as introduction clarity, precision of thesis, use of examples and paragraph unity. No significant improvements were reported for the LOCs of grammar and mechanics. In approach and design, their study was the closest to the present study, although their results could have been affected by permitting students to self-select for the experimental group. Additionally, as the rubric used was not appended and weighting for each component assessed was not reported, it is difficult to determine the range of HOCs and LOCs that were addressed.

In sum, out of ten studies reviewed, all were in the context of tertiary education in the U.S., with no differentiation between participants who were native or non-native speakers of English. Seven of the studies measured impact on overall course grades, rather than directly on writing ability. Two evaluated writing center instruction as an alternative to composition classes, two focused solely on the LOC of students’ grammatical awareness, and six studies possibly introduced the confounding factor of variability through participant self-selection in experimental groups. Finally, none of the studies were situated in an L2 English educational context.

With the growth in writing centers operating outside the U.S. (Rafoth 2014; Hyland 2018), the current study aims to fill a gap in the literature by focusing on the impact of writing center tutorials supplementary to in-class writing instruction in an L2 English higher education context with L1 Arabic speakers. Additionally, it evaluated both the HOCs of writing, such as task fulfillment and organization, as well as the LOCs of vocabulary and grammar. Further, participant self-selection was removed, so variables related to student motivation level were controlled.

4. Aspects of Writing Assessment

In order to evaluate the effect of writing center tuition on students' writing ability, it is necessary to measure student writing ability. Because the writing process is multifaceted, reaching a clearly defined construct of writing ability has been challenging (Slomp 2012). Nevertheless, testing writing proficiency is generally considered necessary, whether it be institution-specific or with commercially-utilized tests, such as the Test of English as a Foreign Language (TOEFL) and the International English Language Testing System (IELTS). Some of the rubrics developed for grading these assessments are holistic, while others are more analytic. The former are more product-oriented, while the latter, (such as the IELTS rubric) assess component parts of a piece of writing, so can lead to more consistent scoring by individual raters. Analytic rubrics are commonly used as an assessment tool in educational programs and with proper training in their use they can improve inter-rater reliability (Rezaei and Lovorn 2010).

Typically, commercial analytical test rubrics break down into component parts which can be categorized as HOCs and LOCs. For instance, the IELTS writing assessment rubric analyzes writing in terms of four equally-weighted aspects, comprising the elements of task response, text organization and cohesiveness, lexical range and accuracy, and grammatical complexity and accuracy. The first two aspects of task response and text organization and cohesiveness correspond to Reigstad and McAndrew (1984, pp. 11–19) HOCs, while the latter aspects of lexical and grammatical range and accuracy relate to LOCs. However, research into the assessment of writing identifies factors that may affect the reliability of raters using such rubrics. These factors include raters focusing unevenly on one or more components of a rubric (Read et al. 2005), rater bias towards handwriting (Klein and Taub 2005), control of basic writing mechanics, and even gender bias (Rezaei and Lovorn 2010). Nevertheless, such rubrics are widely used, with other researchers claiming these rubrics offset rater bias and are hence justified on that basis (Hack 2015; Jonsson 2014). A similar rubric approach was used for the present study.

5. Method

The 65 participants in the study were all female, foundation program students at a university in the Gulf region, who would eventually major in a variety of subject areas. All spoke Arabic as a first language and had varying amounts of English instruction prior to entering the university. However, all participants had completed the same amount of English study, at Level 1 and then Level 2, in the university. Each level comprised 16 weeks of English instruction. At the time of this study, these students were undertaking Level 3 of the English program, which indicated they were at a pre-intermediate level of English ability (IELTS band 4 to 4.5). These students were required to obtain an IELTS band score of five or above before they could commence faculty studies. Level 3 was predominantly focused on preparing the students to take IELTS and had two strands. The integrated strand of nine hours a week focused on vocabulary, grammar, reading and listening, while the writing strand, which comprised four hours a week, was solely involved with aspects of writing. All participants had the same teacher for the integrated strand. Informed participant consent was obtained from students. Students were informed about the study by their English teachers as well as an Arabic-speaking teacher and were given the opportunity to ask questions in their own language. They were given the option to leave the study, at any time, without any impact on their grade point average.

To control for possible variance in teaching pedagogy and content, the two teacher researchers taught all classes involved in the study. Care was taken to ensure that all classes were engaged in identical writing lessons and homework tasks throughout the semester. The researchers were not involved in allocating students to classes. Typically, students in this program chose classes at time slots to fit required courses into their desired timetables. When a class filled, students could no longer enroll in that class. In the first two weeks of semester, students could change classes to resolve timetable clashes. The two teacher researchers each had four classes of approximately 20 students per class within the university foundation program Level 3 writing stream. Consequently, the study began with

eight classes, comprising around 160 students. The two teacher researchers randomly allocated these classes to experimental or control groups before meeting the classes. Each of these researchers had two experimental group classes and two control group classes. The classes were allocated to spread out the load of the student consultations on the Writing Center staff. Therefore, experimental group classes were selected to be at different times on different days for each researcher. However, two weeks into the study, one of the classes was re-assigned for administrative reasons to a different teacher. As a result, one researcher had two experimental group classes and only one control group class. The experimental and control groups then comprised approximately 80 and 60 students respectively.

At the beginning of the semester, a pre-test was administered to all students in the study. This comprised a 250-word opinion essay. The pre-tests were marked using the in-house Level 3 writing strand rubric (attached as Supplementary Material). In this rubric, a total mark out of 20 was divided into four components, with a maximum of five marks possible for each. The components were task fulfillment (TF), organization and coherence (OC), vocabulary (V) and grammar (G). Task fulfillment related to text length, clarity and relevance of ideas, sufficiency of development of ideas and the overall degree to which the essay question was answered. The criteria for organization and coherence related to paragraphing, organization of ideas and text coherence. Criteria for vocabulary evaluated the range and accuracy of vocabulary and spelling, while the criteria for grammar was concerned with the range and accuracy of linguistic structures and punctuation. The principal raters were the two teacher researchers, both of whom were experienced IELTS examiners and also experienced in using the Level 3 writing strand rubric. Inter-rater reliability was developed through two stages of moderation and then analyzed statistically. Firstly, the raters independently marked a sample containing 10 scripts selected as representing a range of levels. Marks awarded by the raters for each script were then compared, discussed and final marks agreed on.

After marking standards had been set in this way, the raters then independently marked a second sample of 30 scripts. These marks were then compared and final marks agreed on for each script. The marks independently given to each of the 30 scripts by the raters were analyzed. The Pearson product-moment correlation coefficient of the two raters was calculated to establish the degree of inter-rater reliability and a strong positive correlation ($r = 0.89$ to 0.96 , $n = 30$, $p < 0.01$) was found, confirming a high degree of interrater reliability. The remaining pre-test writing scripts were then split between the two raters and marked independently. Finally, the marks for the four components as well as the total mark were collated for analysis. Analysis was carried out to establish whether both groups displayed similar writing competencies at the beginning of the study. The mean pre-test scores for both groups were compared. Statistical analysis of overall and component pre-test scores, as detailed in the results section, showed that the control and experimental groups did not exhibit any statistically significant differences in their ability to write an opinion essay at the outset of the study. Therefore, they were well-matched for the purposes of the study in overall scores as well as in the individual component scores.

Throughout the 16 weeks of the semester, both groups completed eight writing portfolio tasks, which contributed up to 3% of their coursework grade for the semester. These portfolio tasks comprised opinion, advantage/disadvantage and problem/solution essays scheduled to fit with the Level 3 writing strand syllabus. Both the control and experimental group participants undertook the portfolio tasks as fluency-focused practice in class and received only very general encouragement and feedback on these eight portfolio tasks from the researchers. However, during the course of the semester, the experimental group participants were required to take each of their eight portfolio pieces to the writing center for one-to-one consultations with a writing tutor. Thus, each experimental group participant was compelled to engage in eight writing center consultations. Students were not obliged to work with the same tutor on each visit. Each consultation lasted approximately 25 min, and during this time, tutors read the draft and gave personalized feedback and advice to the tutee, based on his or her professional analysis of the writing sample. Study participants worked with tutors who were all professional English as a Second Language (ESL) teachers, with a minimum of a master's degree in a

related field and three years' teaching experience. These staff had attended training sessions covering writing center theory and practice, including the use of an interactive, dialogic and non-hierarchical tutoring style. Tutors had been instructed to advocate a process approach to writing and to advise students on aspects of their writing process as necessary. They had been asked to address HOCs in students' work before moving on to LOCs. They had also been asked to read a theoretically-informed and research-based in-house staff handbook (Whitehouse 2013), serving as reinforcement for the training sessions. The tutors conducted the writing consultations in a workspace shared with the writing center coordinator, thus enabling ongoing supervision and support. Tutors were informed of the research study at the beginning of the semester and were asked to sign participants' portfolio booklets on each visit to the center, to indicate both participants' attendance and their satisfactory engagement in the consultation process. Therefore, tutors knew when they were working with a member of the study's experimental group.

The portfolio pieces that the students received feedback on were at the first draft stage, having been written under time constraints in class. The control group students were not required to use the writing center to gain feedback on their portfolio pieces, but they were not discouraged from visiting the writing center for independent consultations.

Throughout the semester, a number of students exited the study. It was usual for around 20% to 30% of Level 3 students to successfully take the IELTS exam externally during a semester and so stop attending English classes. Other students stopped attending English classes for a variety of reasons. Attrition rates during the study were consistent with these norms. This level of participant attrition impacted both experimental and control group numbers, and as such the sample size for each group was reduced.

At the end of the semester, a post-test comprising a problem/solution essay was administered. This was again marked according to the Level 3 writing rubric, using the same protocol followed for developing and statistically confirming interrater reliability as for the pre-tests. The Pearson product-moment correlation coefficient values again showed a strong positive correlation between the marks given by the two raters ($r = 0.90$ to 0.94 , $n = 30$, $p < 0.01$). The marks for the four components as well as the total mark were collated and entered into SPSS by the researchers for analysis. To provide further checking of inter-rater reliability, two colleagues experienced in IELTS and in rating scripts with the Level 3 writing rubric served as secondary raters. Both secondary raters blind rated a sample of the pre-test and post-test scripts from both control and experimental groups. The Pearson product-moment correlation coefficient for all the raters was again calculated and found to be high ($r = 0.86$ to 0.94 , $n = 10$, $p < 0.01$), confirming the validity of the marks given to the participants' pre-test and post-test scripts.

The issue of teacher-effect on students' overall gain scores throughout the semester was also considered. As both researchers taught the control and experimental groups, the likelihood of uneven gain scores impacted upon by possible differences in teaching ability was evaluated. Therefore, overall gain scores of the experimental and control groups were analyzed by teacher group and no significant differences were found. Further analysis was conducted to establish differences in overall gain scores by class. No significant differences were found in overall gain scores among the classes within the control and experimental groups.

Only participants who had completed all eight portfolio tasks were included in the final analysis. As any control group student could have independently used the Writing Center during the course of the study, a survey was conducted to establish which of the control group participants met the criteria for inclusion. Three students from the control group had visited the writing center during the semester and were thus excluded from the study.

6. Results

Thirty participants from the control group were found to have undertaken all eight portfolio tasks and not to have visited the writing center during the semester. Thirty-five participants from

the experimental group were found to have completed all eight portfolio tasks and to have gained feedback from tutors at the writing center for each task. Accordingly, their pre-test and post-test results were analyzed in order to answer the research questions below. Findings are reported in relation to each research question.

1. Did the students in the experimental group and control group significantly improve their writing ability over the 16-week semester?
2. Did the students in the experimental group improve their writing ability as a result of eight consultations at the writing center?

The results for the experimental group are reported in Table 1 below, while results for the control group are shown in Table 2.

Table 1. Experimental group ($n = 35$) overall and component pre-test to post-test mean scores, mean gain scores, and standard deviations (SD).

Measures	Pre-Test Mean	Post-Test Mean	Mean Gain
Overall	8.96 (SD 1.92)	14.50 (SD 1.83)	5.54 (SD 1.68)
Task Fulfillment	1.77 (SD 0.82)	4.00 (SD 0.65)	2.23 (SD 0.77)
Organization & Coherence	2.29 (SD 0.52)	3.93 (SD 0.54)	1.64 (SD 0.59)
Vocabulary	2.41 (SD 0.48)	3.33 (SD 0.51)	0.91 (SD 0.49)
Grammar	2.49 (SD 0.48)	3.24 (SD 0.49)	0.75 (SD 0.35)

Table 2. Control group ($n = 30$) overall and component pre-test to post-test mean scores, mean gain scores, and standard deviations.

Measures	Pre-Test Mean	Post-Test Mean	Mean Gain
Overall	8.90 (SD 1.52)	13.02 (SD 2.27)	4.12 (SD 1.27)
Task Fulfillment	1.78 (SD 0.74)	3.57 (SD 0.64)	1.79 (SD 0.60)
Organization & Coherence	2.33 (SD 0.46)	3.58 (SD 0.53)	1.25 (SD 0.43)
Vocabulary	2.37 (SD 0.41)	3.22 (SD 0.54)	0.85 (SD 0.37)
Grammar	2.42 (SD 0.37)	3.12 (SD 0.60)	0.70 (SD 0.43)

In order to establish if the two groups were equivalent at pre-test, a multivariate analysis of variance (MANOVA) was carried out, comparing the pre-test mean scores of both groups in terms of overall scores and component scores of task fulfillment, organization and coherence, vocabulary, and grammar. The results of the MANOVA showed no significant differences between the groups on any of the measures, $F(4,60) = 0.32$, $p > 0.05$; Wilk's $\Lambda = 0.98$.

The overall scores were analyzed using repeated measures analysis of variance. The results for time showed significant change from pre-test to post-test for both groups, $F(1,63) = 666.03$, $p < 0.01$. The group factor which compared the combined pre- and post-test scores of both groups, was not significant, $F(1,63) = 3.14$. The result of most interest was the significant time \times group interaction which compared the relative improvement from pre to post of both groups, $F(1,63) = 14.54$, $p < 0.01$. It showed that the experimental group had improved more than the control group. An independent samples t -test was used to compare overall mean improvement between the two groups. Results were significant $t(63) = 3.81$, $p < 0.01$ two-tailed. In other words, the improvement of the experimental group, at 5.54 was significantly higher than the improvement of the control group, at 4.12. The difference was 1.42 marks out of a possible 20 marks. Although this seems a small difference, about seven percent, it would be equivalent to more than half a grade. The effect size for this analysis ($d = 0.95$) was found to correspond to a large effect according to Cohen (1988) convention. The result suggests that the students who participated in the writing center sessions made greater gains in their overall writing ability compared to their control group counterparts.

The repeated measures results for the HOC of task fulfillment also showed significant gains in time from pre-test to post-test for both groups, $F(1,63) = 536.79, p < 0.01$. The group effect was not significant, $F(1,63) = 1.82$. The result of most interest was the significant interaction of group and time, $F(1,63) = 6.63, p < 0.05$, which indicated that the experimental group made significantly more progress from pre-test to post-test on the task fulfillment measure than did the control group. The mean gain for the task fulfillment component for the experimental group was 2.23, while for the control group it was 1.78. The difference between the mean gain for the experimental and control groups was 0.45, a nine percent difference. An independent-samples *t*-test showed that the difference in mean gains between the groups was significant $t(63) = 2.58, p < 0.05$, two-tailed. The effect size registered a medium to large effect ($d = 0.65$), indicating that writing center sessions impacted notably on the HOC task fulfillment aspect of students' writing.

Similarly, the repeated measures results for organization and coherence, the other HOC, showed significant improvement from pre-test to post-test for both groups, $F(1,63) = 496.29, p < 0.01$. The group effect was not significant, $F(1,63) = 1.81$. Again, the interaction of group and time was significant, $F(1,63) = 9.15, p < 0.01$, indicating that the experimental group made significantly more progress than did the control group. The experimental group's mean improvement of 1.64 for organization and coherence was 0.39 higher than the control group's mean improvement of 1.25, an eight percent difference. Independent *t*-test analysis showed this difference in relative improvement to be significant, $t(63) = 3.03, p < 0.05$ two-tailed. For this analysis, the writing center sessions had a medium to large effect ($d = 0.76$) on the HOC aspect of organization and coherence of writing.

For the LOC of vocabulary, the repeated measures analysis showed significant improvement from pre-test to post-test for both groups, $F(1,63) = 257.17, p < 0.01$. The group effect was not significant, $F(1,63) = 0.54, p > 0.05$. The interaction of group and time was also not significant, $F(1,63) = 0.34$, indicating that each group made similar levels of improvement. At 0.91 for the experimental group and 0.85 for the control group, the difference between the two groups in terms of relative improvement for vocabulary was 0.06, a very small difference.

The repeated measures analysis for the LOC of grammar showed a significant improvement in grammar from pre-test to post-test for both groups, $F(1,63) = 227.80, p < 0.01$. The group effect was not significant, $F(1,63) = 0.76$. The interaction of group and time was also not significant, $F(1,63) = 0.35, p > 0.05$, indicating that both groups made similar progress. The mean improvement in grammar, an LOC aspect, for the experimental group was slightly higher than for the control group, at 0.75 and 0.70 respectively. The difference in relative improvement, which was 0.05, was a very small difference.

In summary, in relation to research question one, which asked whether both groups improved in writing, both the experimental and control groups improved their scores over the 16-week period. Their overall scores improved as did their subcomponent scores. In relation to question two, which asked whether the experimental group made more improvement than the control group, the results showed that the experimental group improved significantly more in overall scores than did the control group. The experimental group also improved significantly more in two of the subcomponent HOC scores, (a) task fulfillment and (b) organization and coherence, but there was only a small and non significant difference between the two groups in terms of improvement in the LOC aspects of vocabulary and grammar. Possible reasons for these results will be discussed in the following section.

7. Discussion

At the beginning of the study, there were no statistically significant differences between the experimental and control groups in overall writing ability, in their ability to answer a question, organize their writing coherently or write with appropriate range and levels of accuracy of lexis and grammatical structure. By the end of the study, all students, irrespective of group, had improved their writing ability significantly in all the above-mentioned aspects. This is to be expected as all students were participating in a semester-long English program, which included four hours per week of writing instruction and practice. However, students in the experimental group who attended eight writing center consultations

made significantly higher gains in their overall writing scores. The effect size was large. This is consistent with the findings of Niiler (2003, 2005), Henson and Stephenson (2009), and Irvin (2014). Moreover, significant favorable differences were found for HOC aspects of the experimental group's writing, specifically in the areas of task fulfillment and organization and coherence, which support Henson and Stephenson (2009) findings. These were medium to large effect sizes. In contrast, while the experimental group made slightly greater gains in the LOCs of vocabulary and grammar, these were not significantly higher than those of the control group.

The explanation for the greater improvements being made in the HOC areas of task fulfillment and organization and coherence is likely to be found in a combination of factors. The participants brought first drafts to the writing center having spent no or little time on revision in class. Hence, the consultations were the first occasion in which the essays were reviewed in any detail. In keeping with a process approach to writing pedagogy, tutors would have advised students on HOCs, such as content, organization, coherence and cohesion first, and bearing in mind the consultations' 25-min time limit, would only have moved on to lexical and grammatical issues if time permitted. This fits with Flower and Hayes (1981) theory, wherein higher order goals are dealt with early in the writing process. Corresponding to Camp (2012) conception of the writing process as being socially-situated, it is also likely that the tutors' role in giving reader feedback, as described by Harris (1988), helped students gain a raised awareness of the expectations of their target discourse community (the raters of their writing) in relation to textual content and organization. Provisos to be noted here are, firstly, that if students had presented work already well developed in content, argument and organization, the focus of the consultations would have been LOCs (or the consultation would have been curtailed). Secondly, although tutors had been asked to follow the writing center's recommended tutoring approach, there may have been variation in its application between tutors.

The study has shown that working with tutors in the writing center enabled L2 student writers to make measurable improvements in overall writing skills and in writing skills related to HOCs during the course of one semester. Significantly, these were discernible under test conditions, rather than a post-writing (revision) stage. Thus, while working on specific pieces of writing with writing center tutors, L2 writers made gains in transferrable writing skills applicable to the composition of written academic discourse, relating to task analysis, argument development, selection of appropriate content, and textual organization and cohesion, at whole text and paragraph level. It is highly probable that the pedagogical approach and tutoring techniques that define writing center consultations, as outlined by Harris (1988) and Nordlof (2014), accounted for these improvements. The collaborative, non-hierarchical and student-led approach to tutoring (intersubjectivity) means that consultations involve statements and questions (i.e., are dialogic and interactive) aimed at encouraging students to be active in the editing process, review their writing critically, take responsibility for revisions and self-correct when possible. In this way, students have the opportunity to practice processes and skills crucial to their development as autonomous writers, with the guidance of an informed other in a one-to-one personalized context, during which ongoing diagnosis is carried out. Such methods and interactions are likely to aid the internalization of these processes/skills, thereby enabling students to transfer and apply these independently in future writing tasks. Furthermore, it is possible that the consultations had a positive effect on the participants' attitudes to writing and self-efficacy, as found in studies by Davis (1988) and Schmidt and Alexander (2012). This could have led to the participants making improvements in attributes known to correlate with writing ability.

While the tutorial service provided by the writing center to students in the experimental group did not differ from that offered to the general student populace, the conditions under which these students attended are significant. The center is a self-access learning center and as a result students attend voluntarily, meaning they have perceived a need to access help and want to work with a tutor to make progress in specific areas and are sufficiently motivated to do this in their free time. An ability to identify and acknowledge weaknesses, a willingness to seek help and positive motivation levels are factors known to be conducive to learning. However, as the students in the study's experimental group

were compelled to attend the writing center in their free time, it is possible that atypical attitudinal and affective factors came into play during their consultations. Thus, the benefits of the tuition may have been even greater for students who had attended voluntarily compared with the present students who may have shown lower levels of engagement in, or receptivity to the consultation process.

In showing that participation in at least eight writing center consultations over the course of a semester can have a significant, measurable and positive impact on the writing ability of L2 university foundation program students, this study provides tentative support to the case for expansion of writing center work beyond the domains of predominantly L1 English academic communities. It has shown that tutoring practices forged in L1 English academic communities can benefit L2 English student writers in L2 English contexts, and may serve to allay concerns about the appropriateness of transferring writing center practices into new and diverse learning contexts. The study also contributes to the limited bank of empirical studies examining the efficacy of writing center consultations on students' writing ability. Finally, it has shown that quantitative research into writing center effectiveness is a challenging yet achievable goal within the constraints of a university semester.

Perhaps the main limitation of this study was the direct involvement of two of the researchers in teaching the control and experimental groups and rating the pre- and post-tests. A double blind design would have been preferable but there was a lack of funding to engage additional instructors to teach the experimental and control groups, and rate scripts. This direct involvement could have impacted the teaching of the different groups, since the researchers knew which classes were assigned to control or experimental groups, and could also have introduced bias when rating scripts from the different groups. However, the researchers took steps to minimize the effects of this limitation by strictly adhering to the same lesson plans and utilizing the same materials for all the classes. Additionally, to check for bias in rating the control and experimental groups' writing, samples of each were blind marked independently by two other raters and their scores compared with those of the researchers. In order to establish the equivalence of the control and experimental groups at the beginning of the study period, it was necessary to mark all the pre-tests at that time. Therefore, when rating the post-tests towards the end of the study, the raters may have been influenced by expectations of improvement in student writing. However, such expectations of some improvement would have applied equally to the control and experimental groups. It is also worth mentioning that as raters graded handwritten scripts, it is possible that this could have contributed to an additional element of rater variability (Klein and Taub 2005). A further possible limitation was that measurement of the participants' English writing proficiency was operationalized as scores given by raters for one timed essay at pre-test and again at post-test. However, a single sample of writing produced under time constraints may not adequately represent learners' writing ability. Writing performance can vary on different days, introducing an uncontrolled variable into the study. Research using timed essays as a measure of writing ability also provides results that are not necessarily generalizable to writing processes used in tasks carried out without a time constraint. Another possible source of bias were the writing center tutors, who knew when they were working with experimental group participants, which may have had an impact on the consultations. Finally, the findings of this study only apply to female L1 Arabic foundation program students in one university in the Gulf. Therefore, these findings are not necessarily generalizable to all L2 writers.

Further research that utilizes blind marking to control the variable of possible rater bias would help to contextualize the results of this study. Additionally, given that writing centers generally operate within an academic context in which students may more often be required to write coursework assignments rather than undertake single timed writing assessments, future studies could measure the impact of writing center consultations on L2 writers' coursework assignments. This would give further insight into writing centers' relevance to L2 students' progress in writing different kinds of assignments in tertiary level English-medium education settings.

8. Conclusions

In summary, the present study has shown that a total of eight writing center consultations during the 16 weeks of a semester had a large and positive effect on the writing ability of Arabic L1 students in a university foundation program where English was the medium of instruction. From pre-test to post-test, with assessments rated out of a possible total of 20 marks, the overall score of the experimental group was 1.42 marks more than the control group, a seven percent difference that was equivalent to a gain of more than half a grade compared with the control group. Finally, the findings of this research make a case for writing centers to be an integral part of the growth of English-medium education for non-native speakers.

Supplementary Materials: The following are available online at <http://www.mdpi.com/2226-471X/3/1/4/s1>, Level 3 Scoring Rubric—For Project and Exams (Task 2).

Acknowledgments: We appreciate the thoughtful feedback of three reviewers and the editors, which have improved the paper.

Author Contributions: Shanthi Tiruchittampalam was responsible for conception and research design, data collection, and analysis as well as writing Sections 4, 7 and 8 in collaboration with co-authors. Alistair Ross was involved in research design, data acquisition, writing Sections 3, 5 and 7 in collaboration with co-authors, and critical review of drafts for this publication. Elizabeth Whitehouse contributed to the research design and wrote Sections 1, 2 and 7 in collaboration with co-authors. Tom Nicholson contributed to the statistical analysis and write-up of the data from the project in Section 6.

Conflicts of Interest: The authors declare no conflicts of interest.

References

- Babcock, Rebecca Day, and Terese Thonus. 2012. *Researching the Writing Center: Towards an Evidence-Based Practice*. New York: Peter Lang.
- Camp, Heather. 2012. The Psychology of Writing Development and its Implications for Assessment. *Assessing Writing* 17: 92–105. [CrossRef]
- Cohen, Jacob. 1988. *Statistical Power Analysis for the Behavioral Sciences*. Hillsdale: Lawrence Erlbaum Associates.
- David, Carol, and Thomas Bubloz. 1985. Evaluating Students' Achievement in a Writing Center. *Writing Lab Newsletter* 9: 10–14.
- Davis, Kevin. 1988. Improving Students' Writing Attitudes: The Effects of a Writing Center. *Writing Lab Newsletter* 12: 3–5.
- Driscoll, Dana, and Sherry Wynn Perdue. 2012. Theory, Lore and More: An Analysis of RAD Research in the writing center journal, 1980–2009. *The Writing Center Journal* 32: 11–39.
- Eusafzai, Hamid Ali Khan. 2018. The Emergence of Writing Centers in the Saudi Higher Education Landscape: A Genealogical Investigation. In *Writing Centers in the Higher Education Landscape of the Arabian Gulf*. Edited by Osman Z. Barnawi. Basingstoke: Palgrave Macmillan, pp. 3–20. [CrossRef]
- Flower, Linda, and John R. Hayes. 1981. A Cognitive Process Theory of Writing. *College Composition and Communication* 32: 365–87. Available online: <http://www.jstor.org/stable/356600> (accessed on 12 March 2016). [CrossRef]
- Hack, Catherine. 2015. Analytical Rubrics in Higher Education: A Repository of Empirical Data. *British Journal of Educational Technology* 46: 924–27. [CrossRef]
- Harris, Muriel. 1988. SLATE (Support for the Learning and Teaching of English) Statement: The Concept of a Writing Center. Available online: <http://writingcenters.org/writing-center-concept-by-muriel-harris/> (accessed on 10 December 2017).
- Henson, Roberta, and Sharon Stephenson. 2009. Writing Consultations can Effect Quantifiable Change: One Institution's Assessment. *The Writing Lab Newsletter* 33: 1–5.
- Hyland, Ken. 2018. Foreword. In *Writing Centers in the Higher Education Landscape of the Arabian Gulf*. Edited by Osman Z. Barnawi. Basingstoke: Palgrave Macmillan, pp. vii–viii. [CrossRef]
- Irvin, Lennie L. 2014. What a Difference Three Tutoring Sessions Make: Early Reports of Efficacy from a Young Writing Center. *Writing Lab Newsletter* 39: 1–5.

- Jonsson, Anders. 2014. Rubrics as a Way of Providing Transparency in Assessment. *Assessment and Evaluation in Higher Education* 39: 840–52. [CrossRef]
- Klein, Joseph, and David Taub. 2005. The Effect of Variations in Handwriting and Print on Evaluation of Student Essays. *Assessing Writing* 10: 134–48. [CrossRef]
- Lerner, Neal. 1998. Drill Pads, Teaching Machines, and Programmed Texts: Origins of Institutional Technology in Writing Centers. In *Wiring the Writing Center*. Edited by Eric H. Hobson. Logan: Utah State University Press, pp. 119–36.
- Niiler, Luke. 2003. The Numbers Speak: A Pre-test of Writing Center Outcomes Using Statistical Analysis. *Writing Lab Newsletter* 28: 6–9.
- Niiler, Luke. 2005. The Numbers Speak Again: A Continued Statistical Analysis of Writing Center Outcomes. *Writing Lab Newsletter* 29: 13–15.
- Nordlof, John. 2014. Vygotsky, Scaffolding and the Role of Theory in Writing Center Work. *The Writing Center Journal* 34: 45–64.
- North, Stephen M. 1984a. The Idea of a Writing Center. *College English* 46: 433–46. [CrossRef]
- North, Stephen M. 1984b. Writing Center Research: Testing our Assumptions. In *Writing Centers: Theory and Administration*. Edited by Gary A. Olson. Urbana: NCTE, pp. 24–35.
- Puntambekar, Sadhana, and Roland Hübscher. 2005. Tools for Scaffolding Students in a Complex Learning Environment: What have we gained and what have we missed? *Educational Psychologist* 40: 1–12. [CrossRef]
- Rafoth, Ben. 2014. *Multilingual Writers and Writing Centers*. Logan: Utah State University Press.
- Read, Barbara, Becky Francis, and Jocelyn Robson. 2005. Gender, 'Bias', Assessment and Feedback: Analyzing the written assessment of undergraduate history essays. *Assessment & Evaluation in Higher Education* 30: 241–60. [CrossRef]
- Reigstad, Thomas J., and Donald A. McAndrew. 1984. *Training Tutors for Writing Conferences*. Urbana: NCTE.
- Rezaei, Ali Reza, and Michael Lovorn. 2010. Reliability and Validity of Rubrics for Assessment through Writing. *Assessing Writing* 15: 18–39. [CrossRef]
- Roberts, David H. 1988. A Study of Writing Center Effectiveness. *The Writing Center Journal* 9: 53–60. Available online: <http://www.jstor.org/stable/43460772> (accessed on 14 March 2016).
- Sandlon, John A. 1980. The Effects of a Skills Center upon the Writing Improvement of Freshman Composition Students. *Writing Lab Newsletter* 3: 1–3.
- Schmidt, Katherine M., and Joel E. Alexander. 2012. The Empirical Development of an Instrument to Measure Writerly Self-efficacy in Writing Centers. *The Journal of Writing Assessment* 5: 1–10. Available online: <http://www.journalofwritingassessment.org/article.php?article=62> (accessed on 6 November 2017).
- Slomp, David H. 2012. Challenges in Assessing the Development of Writing Ability: Theories, Constructs and Methods. *Assessing Writing* 17: 81–91. [CrossRef]
- Sutton, Doris G., and Daniel S. Arnold. 1974. The Effects of Two Methods of Compensatory Freshman English. *Research in the Teaching of English* 8: 241–49. Available online: <http://www.jstor.org/stable/40170528> (accessed on 14 March 2016).
- Thaiss, Chris. 2012. Origins, Aims, and Uses of Writing Programs Worldwide: Profiles of Academic Writing in Many Places. In *Writing Programs Worldwide: Profiles of Academic Writing in Many Places*. Edited by Chris Thaiss, Gerd Bräuer, Paula Carlino, Lisa Ganobcsik-Williams and Aparna Sinha. Anderson: Parlor Press, pp. 5–22.
- Whitehouse, Elizabeth. 2013. *The Writing Centers' Staff Handbook*, 2nd ed. Al Ain: United Arab Emirates University.
- Wills, Linda Bannister. 1984. Competency Exam Performance and the Writing Lab. *Writing Lab Newsletter* 8: 1–4.

