



Comparative Analysis of the Effects of High-Stakes English Exams on Students' Views of Teachers' Teaching Qualities

Getachew Desalegn

Institute of Foreign Affairs, Ethiopia

<http://dx.doi.org/10.18415/ijmmu.v12i9.7040>

Abstract

High stakes exams might have an impact on test takers' learning objectives. It is unclear how the high-stakes English exams have affected students' evaluations of EFL teachers' teaching qualities, given their diverse academic backgrounds. Thus, the goal of this study was to bridge the knowledge gap by employing a mixed method of embedded research design and collecting data via questionnaire, Focused Group Discussion (FGD), and document analysis. The questionnaire was administered to a stratified random sample of 382, 12th grade students. The students with their one-semester academic achievements were selected from eight secondary schools (Arjo Gudetu, Biftu Gimbi, Biftu Nekemte, Darge, Gimbi, Guliso, Sena Gimbi and Sire) found in East and West Wollega zones, Oromia Regional State, Ethiopia. Data analysis involved descriptive statistics, multivariate and one-way ANOVA. The data gathered through document analysis and FGD were to substantiate the questionnaire. The contents of the five consecutive years of past Ethiopian Secondary School Leaving Certificate English Examination (ESSLCEE) questions were analyzed quantitatively. The qualitative method was used to conduct FGD with the selected student participants of different academic achievement levels from each school. The recorded data was subsequently transcribed, translated, analyzed, and discussed thematically. From the study, it is noteworthy that students with varying academic backgrounds exhibit washback variability. Although there are variations in mean scores between students with varying levels of academic achievement, these differences are not statistically significant when it comes to the students' perceptions of teachers' teaching qualities. That is, the high-stakes ESSLCEE has a greater detrimental impact on the low-achieving groups than it does on the high-achieving ones. The findings imply that there are disparities in students' views about their EFL teachers teaching qualities because of the washback of the high-stakes English examination. The study's conclusions have led to the forwarding of recommendations, and the findings have implications for test takers, educators, testers, curriculum designers and policy makers in Ethiopia and elsewhere.

Keywords: *High-Stakes English Examination; Students' Perceptions; EFL Teachers' Teaching Qualities; Students' Academic Achievements*

1. Introduction

Assessment evolved as the shining star of education, surpassing both teaching and learning. When assessments are used in the teaching and learning process, curriculum designers may receive useful information, educators receive feedback, and learners are differentiated. It is practically hard to discuss teaching and learning in the context of education without assessment or testing (Resnick & Schantz, 2017). As a result, education, learning, and evaluations are trustworthy and integral processes.

Several assessment methods have been used in language instruction, including oral questions, portfolios, paper and pencil assessments, and observation. Even though it is challenging to pinpoint the exact origin of these language assessment tools, the Bible's observation regarding oral pronunciation serves as a superb illustration of one of the first recorded records pertaining to language assessment. According to the Old Testament (Judges 12:5–6) and as demonstrated by Hopkins et al. (1990), the Gileadite army gave the Ephraimites an oral test. The term "shibboleth" is one that the Gileadites required the Ephraimites to speak correctly in order to be rescued, or incorrectly in order to be murdered. These days, testing is one of the primary methods for language assessment that is utilized to

In language education environments, language testing is essential. Language education and learning revolve around it (Islam et al., 2021; Salehi & Yunus, 2012). According to Heaton (1990), working in one of the three areas—teaching, learning, and testing—without being continuously aware of the other is nearly difficult due to the close relationships between them. Language testing is used for a number of purposes, including evaluating educational quality (Cheng, 2005), controlling nepotism in the distribution of limited opportunities (Beikmahdavi, 2016), providing quantifiable figures about the language ability of learners (Cheng & Curtis, 2004), and serving as a criterion for student admission in higher education (Manjarres, 2005).

Language scholars were motivated to conduct research in this field by the idea that testing has a significant impact on teaching and learning, which has been studied since the 19th century (Brown & Bailey, 2008). The name "washback" was coined to describe the phenomena, and it was first made public in 1993 by researchers Alderson and Wall. When comparing the parties involved in high-stakes exams (parents, test makers, teachers, and students), the extent of their effect varies. Compared to other exam stakeholders, the impact on test takers is greater. According to academics, tests and their outcomes have a significant impact on the lives of students who take them directly (Taylor, 2005; Stoneman, 2006; Manjarres, 2005). Furthermore, a number of research have documented the impact of high-stakes public examinations on students' learning (Friska & Setiawan, 20018; Nahdia & Trisanti, 2019; Kutlu et al., 2020; Rahman et al., 2021). Nevertheless, it was forgotten how the high-stakes exam affected the test-takers' views towards their EFL teachers' teaching qualities (Buyukkeles, 2016). Furthermore, the impact of high-stakes public exams remains unexplored, as they may have unequal effects on participants from different backgrounds (Cheng, 2005; Pan, 2014). There hasn't been any research done on the dual problems in relation to the current Ethiopian Secondary School Leaving Certificate English Examination (ESSLCEE) setting (Gashaye, 2021 & 2020; Leta, 1990; Negede, 2002). Therefore, the current study looks into how ESSLCEE affects students' perceptions of EFL teachers' instructional styles as well as how different exam results affect applicants based on their academic standing. The following research questions are developed to close the indicated gap:

1. To explore the extent of the influence of ESSLCEE on students' perceptions of effective English teachers' teaching quality.
2. To determine if there is statistically significant difference in the students' perceptions about effective English teachers' teaching qualities because of their academic achievements.

2. Literature Review

2.1 Language Teaching-Learning, Testing and Washback

Testing has an impact on the process of teaching and learning. It is a powerful instrument that looks at the lessons that teachers are giving their pupils (Tayeb et al., 2014). Testing is viewed as the primary instrument to gauge the achievement of the desired goal, with learning being regarded as the essential goal of instruction (Saif, 2006). Consequently, testing has an unavoidable and consequential impact on the teaching and learning process since teaching, learning, and testing are interdependent components (Shih, 2006). High-stakes public exams, in particular, have an impact on how teaching and learning are implemented in classrooms (Spratt, 2005; Rea-Dickins & Scott, 2007; Al-Hinai & Al-Jardani, 2020). For instance, Spratt (2005), in his study, revealed that tests impacted teaching-learning narrowing the curriculum by more focusing on the examinable areas, but skipping the remaining course contents. The term "washback" refers to the overall impact of testing on education (Alderson & Wall, 1993; Bailey, 1996). For example, Spratt's (2005) study found that exams had an impact on teaching and learning by making the curriculum more narrowly focused by emphasizing the parts that could be tested and ignoring the other course material. The term "washback" refers to the overall impact of testing on education (Alderson & Wall, 1993; Bailey, 1996).

Testing has a positive or bad impact on teaching and learning. Considerable care must be used in determining how it might help or impede the achievement of teaching-learning objectives. Positive washback of a given test occurs when teaching and learning are limited to their intended purposes, which are the same as preparing for that test (Taylor, 2005). On the other hand, a test is considered to have negative washback if there is a mismatch between the intended teaching-learning purposes and the participants' focus during the test, leading them to disregard the course's intended purposes in favor of the test format (Brown & Bailey, 2008).

Washback research is becoming more complicated and dynamic in nature since it incorporates a range of intervening variables, such as testing, and its stakeholders (parents, teachers, school administrators, and curriculum designers) (Alderson & Wall, 1993; Cheng, 2008; Green, 2007). It was unable to draw a clear distinction between the impacts of examinations on instructional strategies in the classroom (Alderson & Wall, 1993). It is mediated by cumulative influences and can be characterized by a range of attributes. There are several factors that work together to potentially influence how testing affects teaching and learning (Green, 2007). The scenario makes the effect complex because there are several factors that contribute to the occurrence of the washback effect of testing on the teaching-learning process.

It is undeniable that high-stakes tests have an impact on test-takers' learning behaviors (Washback, 2013). The problem has a special relationship to students among those involved in the impact of testing on teaching and learning (Taylor, 2005; Stoneman, 2006; Manjarres, 2005). The influence of high-stakes exams is primarily beneficial to test takers. Given that tests have a direct impact on students' lives, it is agreed that testing has the capacity to generally have an impact on those lives (Cheng, 2008; Spratt, 2005). Accordingly, even if a number of research (Alderson & Wall, 1993; Bachman & Palmer, 2010; Bailey, 1996; Cheng, 2005; Green, 2007; McNamara, 2000) were carried out, it is believed that they were less in number than other significant participants, teachers. However, more research is still required to determine how the high-stakes ESSLCEE affects students' opinions of EFL teachers' instructional styles and to characterize any disparities that may exist amongst students based on their academic standing.

2.2 Context of the Present Study: Historical Background

Ethiopia has been doing national exams for a very long time. The London General Certificate of Education test was the name of the first national test for secondary schools (Gashaw, 1982; Leta, 1990). The test was created in the United Kingdom and brought to Ethiopia to be administered to grade 12 pupils. Addis Ababa University started administering the "Ethiopian School Leaving Certificate Examination (ESLCE)" in Ethiopia in 1950 (Gashaw, 1982). Students who had finished grade ten were given the opportunity to take the Ethiopian General Secondary Education Certificate Examination (EGSECE), a new national examination program that was implemented in 2001 (Mulu et al., 2009). The test was used to filter applicants before they were accepted into colleges or preparatory institutions for vocational training. In 2003, the Ethiopian Higher Education Entrance Examination (EHEEE) took the place of the ESLCE, which had been superseded by grade ten national examinations. Later, starting in 2019, the Ethiopian Secondary School Leaving Certificate Examination (ESSLCE) replaced the Ethiopian Higher Education Examination (EHEEE), which had been eliminated from the educational system (Mulu et al., 2009). The Ethiopian National Educational Assessment and Examination Agency is currently in charge of overseeing the preparation of the ESSLCE (ENEAEA).

In Ethiopia, teaching English as a foreign language is required beginning in primary school (grades 1–8) and continuing through secondary school (grades 9–12) and higher education (FDRGE, 1994; Gerenchea, 2019). Students take the national exam after completing their schooling up to grade 12. The candidates' potential national examination score is very important for their future careers and higher education. English is one of the subjects that students take on the national test in addition to other disciplines. Since the English subject's score on the national exam, together with the scores of the other subjects, has the potential to significantly impact students' future lives, it can be considered a high-stakes exam (Stoneman, 2006; Manjarres, 2005).

2.3 Empirical Studies

Testing's impact on teaching and learning, or WASCHCK, is intricate and multifaceted. It is difficult to define exactly what washback is these days, having begun with the 1990s when it was declared to exist (Alderson & Wall, 1993). The possible lack of a linear relationship between assessments and the teaching-learning processes is the cause (Shih, 2006). Through the use of models, researchers have attempted to reveal the intricate nature of the washback (Alderson & Wall, 1993; Hughes, 2003; Rea-Dickins & Scott, 2007). According to Hughes (2003), an exam can have a beneficial or bad effect on the teaching and learning processes of its participants, including teachers, students, and material designers.

It is imperative that significant attention be paid to students, who are among the main players in the influence of testing. According to certain research, it's crucial to look into how EFL students learn and how high-stakes tests affect them (Alderson & Wall, 1993; Brown & Bailey, 2008; Green, 2007; McNamara, 2000; Gashaw, 1982). Examined in the context of grade 12 ESSLCEE is the impact of testing on students' learning, which has received special emphasis in this study.

According to a review of the relevant evidence, high-stakes exams can either positively or negatively affect EFL students' learning (Taylor, 2005; Brown & Bailey, 2008). Conversely, due to individual variances in test-taking, applicants may experience differing washback effects (Alderson & Wall, 1993; Stoneman, 2006; Pan, 2014). For instance, Alderson and Wall (1993) suggested in their theory that different learners may have distinct washback from a test. In a similar vein, it is noted that test washback variance is seen among various learner groups based on their study years, competence levels, and test perceptions (Pan, 2014).

Studies on high-stakes exams and their impact on students' learning tendencies have been conducted, but less attention is paid to how the testing affected the participants' EFL teachers' abilities as

teachers. Furthermore, in the current study's context, the impact of EFL testing on students' individual differences—for example, due to differences in their academic achievement—remains unaddressed. Therefore, the study's main focus is on how high-stakes public exams affect test-takers' perceptions of EFL teachers' instructional styles and how their perspectives alter based on academic performance.

3. Research Method

A mixed-method causal comparative design was used to ask the aforementioned research questions. Such a design is thought to aid in the comparative description of the attitudes, beliefs, and traits of the sample population (Creswell & Creswell, 2018). It is primarily selected to ascertain the opinions of individual students concerning the impact of ESSLCEE on students' perceptions of their own learning. The study prioritized the research issues to be addressed sequentially using both quantitative and qualitative methodologies. Eight public secondary schools located in the East and West Wollega zones—Arjo Gudetu, Biftu Gimbi, Biftu Nekemte, Darge Nekemte, Gimbi, Gulliso, Sena Gimbi, and Sire—are the study sites. A random selection of 382 grade 12 students who were enrolled in classes at both institutions during the data collecting period made up the sample for the questionnaire. Six students from each school took part in the focus groups. Students who were willing to share their rich and varied perspectives into the issue under inquiry, as well as their differences in academic achievement, were taken into consideration when choosing the discussant students (Dörnyei, 2007).

One method used to collect data and draw conclusions regarding respondents' self-reported practices was the questionnaire (Punch, 2005; Bryman, 2008). FGDs with students and papers (the last five years' worth of ESSLCEE questions) were examined in order to support the information gathered via the questionnaire. Prior to the first round of data collecting in November 2022, the study sites were visited in order to secure administrative body permission. Following that, the data collecting was done in a methodical manner. First, a survey was sent out in March 2022, and in April 2022, focus group discussions took place. Finally, in June 2022, examples of previous ESSLCEE questions as well as the academic standing of the students were gathered from the chosen institutions.

Following the collecting of data, the research questions guided the organization of the findings' presentation and conversations. While the results from the focus group discussions (FGD) were examined subjectively, the data from the questionnaire and samples of previous ESSLCEE publications were studied objectively. In order to address the first research question, the analyses and conclusions from the student questionnaire, along with the FGD and previous ESSLCEE questions, were presented concurrently. Data from the questionnaire was triangulated with information from the students' FGD and the document analysis of previous ESSLCEE questions. Students' academic success records and self-report data were used in a one-way ANOVA to address research question number two, one of the inferential statistics. Subsequently, the outcomes derived from both quantitative and qualitative data were discussed, summarized, and concluded.

Reviewers, who possess PhDs in linguistics and TEFL, verified the validity of the items in the FGD and student questionnaire. Cronbach's alpha was used to assess the reliability of the items' internal consistency. Accordingly, students' judgments of teachers' instructional qualities in relation to their EFL perspectives have a satisfactory items reliability of 0.812 (Pallant, 2020).

4. Results

4.1 General Overview

In this first section of the report, the impact of high-stakes ESSLCEE on the students' perceptions of effective English teachers' teaching qualities is analyzed using group comparisons. The items used for data collection were modified from the questionnaires used in the past studies of the qualities of effective English teachers as perceived by learners (Park and Lee, 2006; Zurrahmi and Triastuti, 2022). During the analysis, the data collected using the questionnaire was categorized into two: those linked to the ESSLCEE and those not related to the ESSLCEE. The understanding constructed from a comparative analysis of the two data types was made to address the research question. Figure 1 below summarizes the data obtained from the ESSLCEE- related items.

4.2 The Influence of the ESSLCEE on Students' Perceptions of Effective English Teachers' Teaching Qualities

4.2.1 ESSLCEE-Related Items

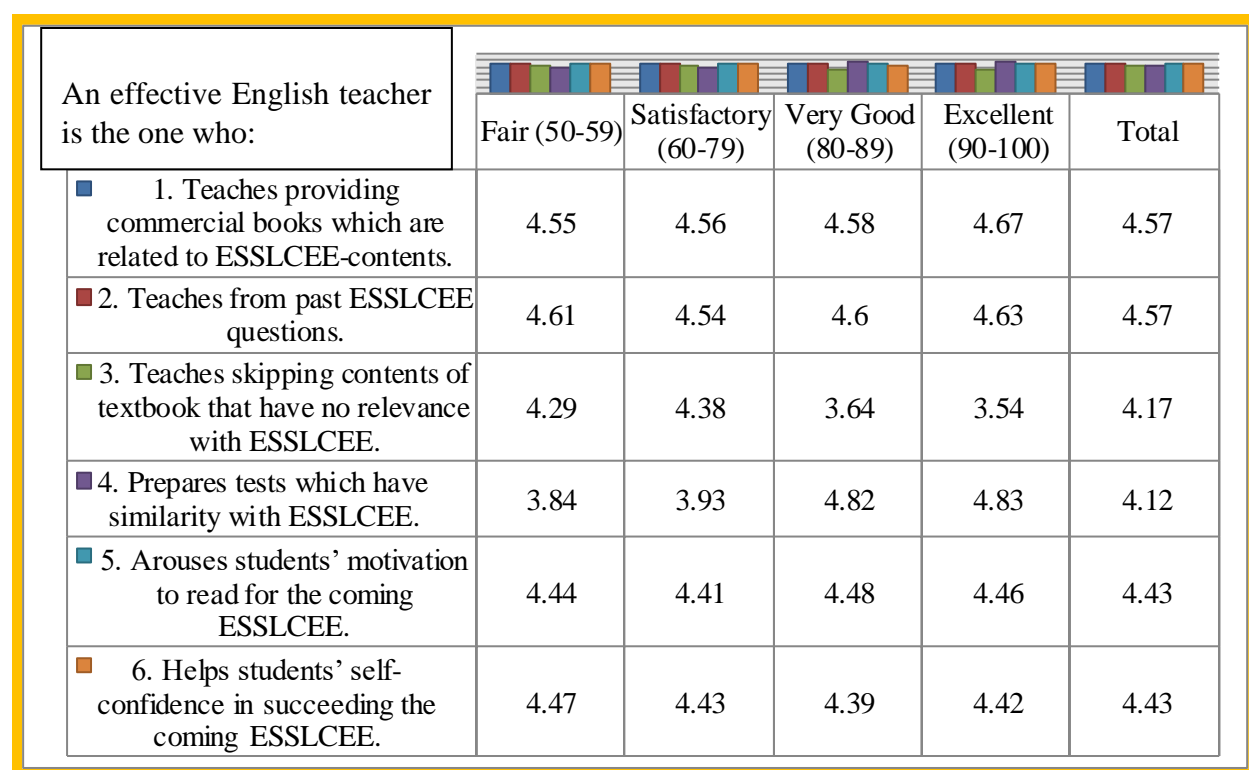


Figure 1. Responses to influence of ESSLCEE on students' perceptions of effective English teachers' teaching qualities (ESSLCEE-related items)

A comparison of the total mean scores registered in terms of the six ESSLCEE-related items, as shown in Figure 1, reveals that $M = 4.57$ is the highest mean score documented in the displayed data. This mean score was obtained in response to the 2nd item in Figure 1. The item was designed to know the students' opinion of an English teacher who teaches them from the ESSLCEE-related commercial books and the past ESSLCEE questions. The students' academic achievement level has not been observed influencing the students' perceptions.

On the other hand, the total mean score $M = 4.12$; $SD = 0.623$ is the lowest mean score documented in Figure 1. This low mean score was obtained in response to the 30th item in Figure 1. The item asked the respondents about their opinion of a teacher who prepares tests which are similar with the ESSLCEE questions. In connection with the students' responses to item 3 - a teacher who skips non-ESSLCEE-related textbook contents - and item 4 - a teacher who prepares classroom tests that are similar with the ESSLCEE questions, noticeable mean differences have been documented among the achieving groups. For example, regarding item 30, i.e., the *Fair* achievers registered a mean value of $M = 3.84$; $SD = 0.479$. The *Very Good* achievers registered the mean value of $M = 4.88$; $SD = 0.327$. Similarly, *Satisfactory* achievers registered a mean value of $M = 4.38$; $SD = 0.661$ in response to item 3, the *Excellent* achieving students, on their part, registered the mean value of $M = 3.54$; $SD = 0.509$. This shows differences in the perceptions among different achieving groups. The data shows that an English teacher who skips teaching non-ESSLCEE-related contents from the textbook (item 3) is preferred by *Fair* and *Satisfactory* achieving groups. Both are low achieving groups. More students from the *Very Good* and *Excellent* achieving groups chose a teacher who prepares classroom tests which have similarity with the ESSLCEE questions than did the students from the *Fair* and *Satisfactory* groups.

The FGD data confirms the questionnaire data collected in this regard. For example, one of the *Very Good* achieving students explained his views about the qualities that distinguish effective English teachers from those who are not as follows:

In my understanding, a teacher who teaches from the commercial books that have relevance with the ESSLCEE questions is a clever teacher. This year, I am interested in my English teacher because of this.

Another participant from the *Fair* achieving group expressed his view which has a connection with the data generated through item 4.

In my opinion, an effective English teacher is not one who prepares classroom tests which appear on the ESSLCEE examination, but the one who explains the questions and let the students prepare for the examination. The inclusion of questions which may appear in the ESSLCEE in classroom tests reduces the students' score on their transcript, the student added.

One of the *Satisfactory* achievers reported his responses to items 2 and 3 saying that a teacher who restricts himself to using the textbook contents and who does not bring the past ESSLCEE questions to the classroom for teaching is not the right teacher.

Among the *Excellent* achievers, a participant in the FGD characterized an effective teacher of English as follows:

I think an effective English teacher is not the teacher teaches ignoring the textbook's contents which did not appear in the past ESSLCEE questions. For me, effective English teacher is the teacher who teaches all the lessons in the textbook to improve the students' communicative skills. Such a teacher relates the past ESSLCEE questions with the textbooks' contents. I am happy with what our current English teacher is teaching us. The reason is that he engages us in the practice activities in the textbook and in other materials. He also teaches us from the past ESSLCEE questions, relating them with the textbook contents. He also prepares classroom tests that are similar to the ESSLCEE questions. So, I am interested in the type of a teacher who restricts himself to textbook contents and does not bring past ESSLCEE questions to class for teaching.

The analysis of the FGD data and that of the quantitative data suggests the existence of variations in the students' perceptions of their English teachers' qualities of instruction. With the exception of the data obtained in reply to items 27 and 30, replies to all the other ESSLCEE-related items (See Figure 1) have high total mean scores. This is true regardless of the respondents' academic attainment levels, compared to the data obtained in response to the non-ESSLCEE-related questions (Figure 2). The observed mean differences between the items connected to the ESSLCEE and those not related to the ESSLCEE have implications for how the high stakes ESSLCEE examination will affect the students' perception of the instructional qualities of effective English teachers.

4.2.2 Non-ESSLCEE-Related Items

The basic purpose of incorporating non-ESSLCEE-related items into the questionnaire is to examine the extent to which the students prefer the items, compared against the ESSLCEE-related items. The result may help to draw conclusions about the students' perceptions of the influence of high-stakes ESSLCEE on the quality of English teachers' teaching.

The non-ESSLCEE-related items used in this study were adapted from Park and Lee (2006) and Zurrahmi and Triastuti (2022). As shown in Figure 12, the non-ESSLCEE-related items fall into three categories: socio-affective skills (items 15, 16, 17; pedagogical skills (items 10, 11, 12, 13, and 14 and subject matter knowledge (items 7-9).

The items on the socio-affective skills achieved the greatest mean score of all the groups. This happened regardless of the students' academic achievement levels. Responses to the pedagogical skills items, on the other hand, had the lowest mean scores. In addition, there is a considerable variation in mean scores across the four achieving groups for the pedagogical skill items when compared to the mean scores registered for subject matter knowledge and socio-affective skills. For instance, in response to a teacher who teaches English in English, the *Very Good* and the *Excellent* achieving groups recorded $M = 3.27$; $SD = 0.880$ and $M = 3.25$; $SD = 0.897$ respectively, while the *Fair* and the *Satisfactory* achieving groups recorded $M = 2.59$; $SD = 0.494$ and $M = 2.69$; $SD = 0.771$. The data manifests notable mean differences between the two categories of the four achieving groups: *Fair* and *Satisfactory* achieving groups in one category and *Very Good* and *Excellent* achieving groups in the other category. Figure 2 below summarizes the data on the influence of the ESSLCEE on the students' perceptions of effective teachers' teaching qualities.

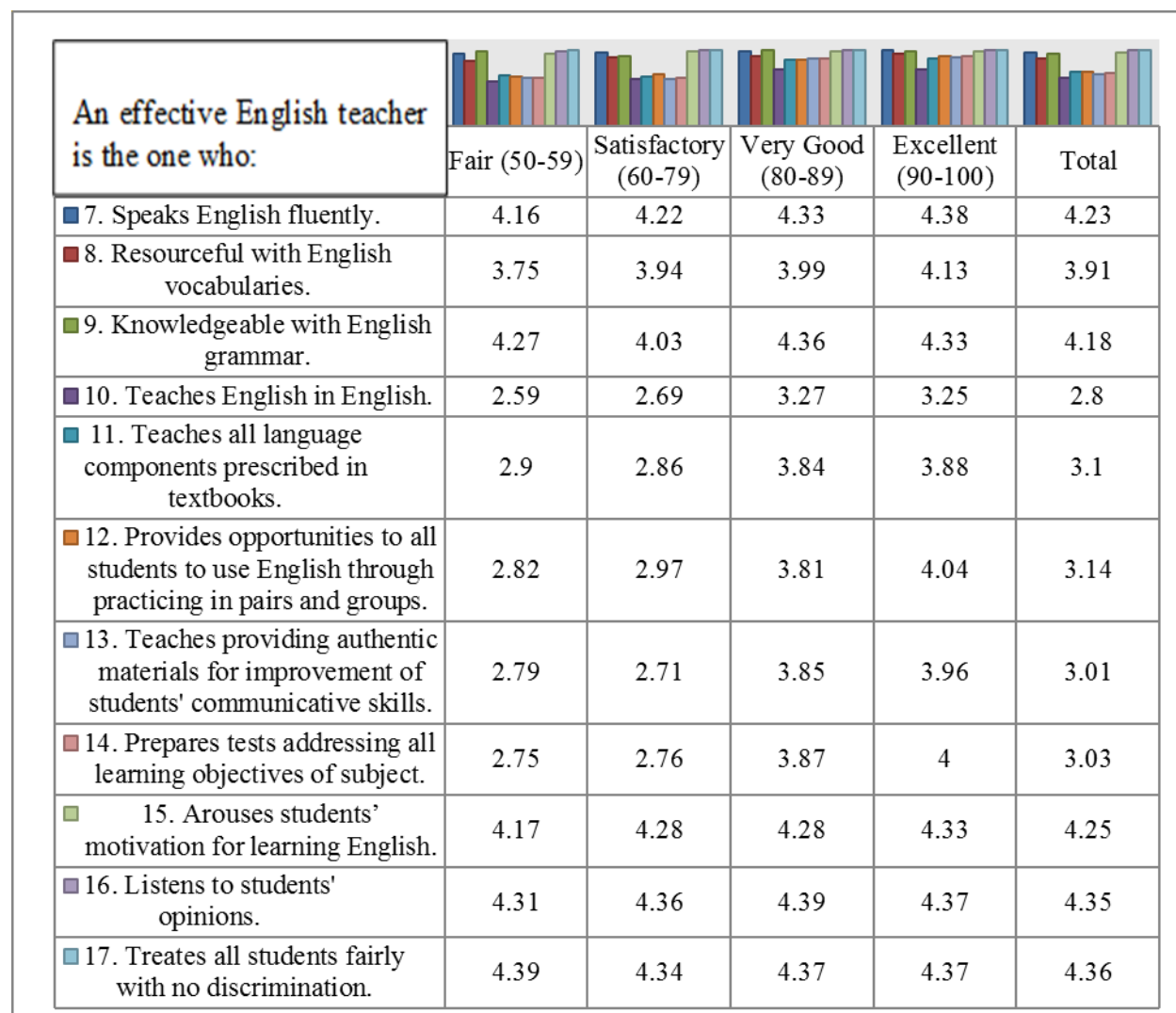


Figure 2. Responses to influence of ESSLCEE on students' perceptions of teachers' teaching qualities (non-ESSLCEE-related items)

The high achieving groups, i.e., *Very Good* and *Excellent* achieving groups, acknowledged an English teacher who teaches English in English, teaches all the language components presented in the textbooks, engages students in meaningful classroom interactions in pairs and groups, provides authentic materials to improve the students' communicative skills, prepares classroom tests targeting the attainment of the objectives of the subject more than did the low achieving groups, i.e., *Fair* and *Satisfactory* achieving groups. However, regarding the mean scores across the four achieving groups for subject matter knowledge and socio-affective skills, high and closer mean scores are registered. For example, in response to a teacher who arouses students' motivation for learning English *Very Good* and *Excellent* achieving groups documented $M = 4.28$; $SD = 0.454$ and $M = 4.33$; $SD = 0.482$ respectively, while the *Fair* and *Satisfactory* achieving groups recorded $M = 4.17$; $SD = 0.425$ and $M = 4.28$; $SD = 0.561$. Thus, the variations of mean scores which exist among the four achieving groups are not notably observed as that of the mean scores of the items under the category of the pedagogical skills. The reports from FGDs corroborate with the data from the questionnaire.

In connection with the issues of subject matter knowledge and socio-affective skills, responses of different achieving groups are modeled below. The responses reveal the similarities in the participants' perceptions. Regarding the subject matter knowledge, one of the *satisfactory* participants mentioned that a

teacher who teaches English should be better than his students in their knowledge of the subject. Another respondent from the *Very Good* achieving group reported that an effective English language teacher is the one who masters all the skills of the language they are teaching. Concerning the socio-affective skills, one of the *Excellent* achievers held the view that ‘*an English teacher, as well as teachers of the other subjects, should not make any bias against their students based on their kinship or religion*’. Similarly, a *Fair* achiever stated that an English teacher should understand the feelings of their students and welcome their students’ opinions accordingly.

The perceptual differences observed between *Fair* and *Satisfactory*, and *Very Good* and *Excellent* achieving groups are also substantiated further by the FGD data. For instance, one of the *Very Good* achiever participants explained his views about the requirements that describe an English teacher as effective in the following way:

As I think, the teacher should help us to practice all the activities in the textbook and even from additional material to enable us to use the language effectively. We are not only preparing for examinations; we are learning the English language skills because they are the bases of our further education. As I evaluated what and how the current English teacher is teaching us, it is not good. We are not satisfactorily practicing the activities in the textbook.

Another member of the *Excellent* achieving group indicated that learning English was important to him for reasons other than just passing the national examination. He is willing to employ the language in his future scholarly work. Extending his justifications, he labeled an English teacher as a bad teacher if they don't assist students in using English. The participant's explanation is shown below in both its original and translated forms:

Personally, I learn English to use the language effectively. I am not learning English only to pass entrance exam. If I am poor in English, in university, I could not score good grade in the courses I take. If my English teacher does not engage me with activities that help me to use the language, he is not the right teacher.

Another view which is almost opposite to the *Very Good* and *Excellent* achievers’ view is forwarded by one of the *Satisfactory* achievers as such:

As to me, as we are preparing for national examination, a teacher should not attempt to teach all activities found in the textbook. In my class, our current English teacher is teaching us considering our interests.

In a nutshell, the impact of high-stakes ESSLCEE on students' perceptions of English teachers' teaching qualities is visible when the data obtained from non-ESSLCEE-related and ESSLCEE-related items are compared. The low achieving groups perceived their English teachers' teaching qualities more negatively than did the high-achieving groups. This finding agrees with the findings reported by Owusu (2021), Park and Lee (2006) and Wichadee (2010).

To determine whether or not significant differences exist among the four achieving groups, the mean scores were computed using one-way ANOVA (see Table 1).

Table 1. Significance test results for students' views towards EFL teachers' teaching

		Sum of Squares	df	Mean Square	F	Sig.
Non-ESSLCEE-Related Items ('number 19-22, 23, 25, 28, 29, 31, 33 and 34')	Between Groups	2088.662	3	696.221	158.318	.000
	Within Groups	1662.293	378	4.398		
	Total	3750.955	381			
ESSLCEE-Related Items ('number 24, 26, 27, 30, 32 and 35')	Between Groups	6.163	3	2.054	.900	.441
	Within Groups	862.999	378	2.283		
	Total	869.162	381			

As shown in Table 1, the test result for the mean score of responses given in response to ESSLCEE-related inquiries about students' opinions of their EFL teachers' instruction is statistically insignificant. It shows that there were no discernible differences in how students perceived the instructional qualities of EFL teachers that were relevant to the ESSLCEE questions among students of various academic achievement levels. However, the test result of the students' perceptions of the non-ESSLCEE-related items that show their EFL teachers teaching characteristics demonstrate as a test of statistical result is significant ($p < 0.05$). The specifics of the test are shown in Table 2 below.

Table 2 . Post Hoc Tests (Tukey HSD) students' views towards EFL teachers' teaching of non-ESSLCEE-related items

(I) Achievment	(J) Achievement	Mean D/ce (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Fair (50-59)	Satisfactory (60-79)	-.237	.254	.787	-.89	.42
	Very Good (80-89)	-5.440*	.325	.000	-6.28	-4.60
	Excellent (90-100)	-6.123*	.472	.000	-7.34	-4.90
Satisfactory (60-79)	Fair (50-59)	.237	.254	.787	-.42	.89
	Very Good (80-89)	-5.204*	.300	.000	-5.98	-4.43
	Excellent (90-100)	-5.887*	.456	.000	-7.06	-4.71
Very Good (80-89)	Fair (50-59)	5.440*	.325	.000	4.60	6.28
	Satisfactory (60-79)	5.204*	.300	.000	4.43	5.98
	Excellent (90-100)	-.683	.499	.519	-1.97	.60
Excellent (90-100)	Fair (50-59)	6.123*	.472	.000	4.90	7.34
	Satisfactory (60-79)	5.887*	.456	.000	4.71	7.06
	Very Good (80-89)	.683	.499	.519	-.60	1.97

*The mean difference is significant at the 0.05 level.

From the findings shown in Table 13, it can be deduced that *Fair* achievers perceived non-ESSLCEE-related items considerably differently from *Very Good* and *Excellent* achievers. The difference between the *Satisfactory* achieving group and the *Very Good* and *Excellent* ones is also quite large. However, there is little difference between the *Fair* and *Satisfactory* ratings. The *Very Good* and the *Excellent* achieving groups show a similar lack of a statistically significant difference. The results indicate that low achievers were less engaged with a teacher who did not focus on language skills that would help them do well on high-stakes exams. The statistics show that high achieving groups had a more favorable opinion of the EFL teachers' teaching qualities regarding non-ESSLCEE-related items than did the low achieving groups, i.e., *Fair* and *Satisfactory*. The findings reported by Owusu (2021), Park and Lee (2006), and Wichadee (2010) are somewhat relevant to the current findings.

4. Discussion

Analysis that has been made under is about the influence of high-stakes ESSLCEE on students' perceptions of English teachers' teaching qualities. Comparing the data obtained from non-ESSLCEE-related and ESSLCEE-related items, the impact of high-stakes ESSLCEE on students' perceptions of English teachers' teaching qualities is noticed. Regardless of the students' academic achievement levels, the results from questionnaire and FGD show that students more favored an English teachers who teach specific contents or materials and strategies that help them to succeed in the coming ESSLCEE than teachers those strive to help them improve the general English. The test result of students' perceptions of English teachers' teaching qualities regarding the non-ESSLCEE-related items shows as statistically significant difference is reported among students of different achieving groups. Regarding the test result of ESSLCEE-related inquiries, however, statistically insignificant difference is reported. The results infer that high achieving groups more acknowledged an EFL teachers who teach English in English, teach all language components prescribed in the textbooks, engage students in meaningful classroom interactions by involving them in pairs and groups, provide authentic materials which improves students' communicative skills and prepare classroom tests targeting the attainment of the objectives of the subject when compared to the low achieving groups (see Figure 1). It is examined that the impact of high-stakes ESSLCEE on students' perceptions of EFL teachers' teaching qualities is remarkable when the data obtained from non-ESSLCEE-related items and ESSLCEE-related items are compared. It is reported that students of the low achieving groups perceived their English teachers' teaching qualities more negatively as a result of the high-stakes ESSLCEE. As far as the existing literature on washback studies is reviewed, the current finding has brought new insight about the influence of high-stakes examination on students' perceptions of teachers' teaching qualities. Studies conducted by Owusu (2021), Park and Lee (2006) and Wichadee (2010) which were conducted purely on teachers' teaching qualities as perceived by the students have some kind of connection with the present finding. Owusu's study revealed that students wanted their teachers to concentrate on language areas that would only make them perform well in the high-stakes tests. Park and Lee investigated the qualities of effective English teachers as perceived by high school students in Korea. The result of the study shows that variables such as gender and students' academic performance seemed to have contributed significantly in the way students ranked certain qualities of effective English teachers. Similarly, a study conducted by Wichadee at Bangkok University, Thailand revealed that statistically significant differences were found in the students' perceptions among different levels of English proficiency towards the qualities of effective English language teachers. Like the current study, however, Owusu, Park and Lee's and Wichadee's studies did not give attention to the perceptual differences may exist among students of different achieving groups because of the influence of high-stakes English tests on students' perceptions of English teachers' teaching qualities.

5. Conclusion

In general, conclusions were drawn based on the findings of the study. As such, the influence of ESSLCEE on students' perceptions towards EFL teachers' teaching qualities was clearly noticed. Particular to the non-ESSLCEE related responses, there is significant differences among students because of their academic achievement levels. The study was carefully designed, data was ethically collected and the results were critically analyzed. However, there are some limitations that need to be considered for the future research to be conducted in the area. First, since such study has not been conducted in the context of the country so far, there might be some more questions and FGD items needed to be included to obtain more data in depth. Second, the study focused on few selective schools which may not exactly represent the remaining schools found in the country from corner to corner. Hence, future researchers can incorporate more items of the questionnaire and FGD, and extend their study areas where the current study excluded.

Disclosure Statement

The author has no competing interests.

References

- Alderson, J. C., & Wall, D. (1993). Does washback exist? *Applied Linguistics*, 14(2), 115-129.
- Al-Hinai, M., & Al-Jardani, K. (2020). Washback in language testing: An exploration with a focus on a specific EFL context in Oman. *International Journal of English Linguistics*, 11(1), 68-75.
- Bachman, L.F., & Palmer, A. (2010). *Language assessment in practice: Developing language assessments and justifying their use in the real world*. OUP.
- Bailey, K. M. (1996). Working for washback: A review of the washback concept in language testing. *Language testing*, 13(3), 257-279.
- Beikmahdavi, N. (2016). Washback in language testing: Review of related literature first. *International Journal of Modern Language Teaching and Learning*, 1(4), 130-136.
- Brown, J. D., & Bailey, K. M. (2008). Language testing courses: What are they in 2007? *Language Testing*, 25(3), 349-383.
- Bryman, A. (2008). *Social Research Methods* (4th ed.). OUP.
- Buyukkeles, G. (2016). *The Washback Effect of a High-Stakes Exit Test on Students' Motivation in Turkish Pre-University EFL Preparatory School* [Master's thesis, University of Reading].
- https://www.teachingenglish.org.uk/sites/teacheng/files/mda2017_university_of_reading_guliz_buyukkelles.pdf
- Cheng, L. (2005). *Changing language teaching through language testing: A Washback study*. Cambridge: CUP.
- Cheng, L. (2008). The key to success: English language testing in China. *Language Testing*, 25, 15-37.
- Cheng, L., & Curtis, A. (2004). *Washback or backwash: A review of the impact of testing on teaching and learning: Washback in language testing: Research contexts and methods*. Lawrence Erlbaum Associates.
- Creswell, J.W., & Creswell, J.D. (2018). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches* (5th ed.). Thousand Oaks: Sage.
- David J. (2000). *Handbook of Parametric and Nonparametric Statistical Procedures* (2nd ed.). Chapman & Hall.
- Degefu, D. (2017). *The Washback Effect of Ethiopian General Secondary Education Certificate English Examination (Egsecee) On Students' Learning Preference: The Case of Shirka Secondary School* [Unpublished master's thesis]. Adama Science and Technology University.

- Dörnyei, Z. (2007). *Research methods in applied linguistics: quantitative, qualitative, and mixed methodologies*. Oxford University Press.
- Friska, Y., & Setiawan, I. (2018). Students' Perceptions towards National Examination: A Washback Study. *Academic Journal PERSPECTIVE: Language, Education and Literature*, 6 (2), 144-158.
- Gashaw, T. (1982). *The Measurement of English Language Proficiency of High School Graduates* [Master's thesis, Addis Ababa University]. AAU Institutional Repository. <http://etd.aau.edu.et/handle/123456789/8747>.
- Gashaye, S (2021). **The Washback of the Ethiopian National English Examination (ENEE) for Grade 10 on Students' Practice:** Debre Markos Secondary Schools on Focus. *Theory and Practice in Language Studies*, 11 (7), 780-787.
- _____ (2020). The Washback of English National Examination for Grades 9 and 10 on Teaching Practices in Ethiopia: A Case Study of Debre Markos Secondary Schools. *International Journal of English Language and Literature Studies*, 9 (2), 106-120.
- Green, A. (2007). IELTS washback in context. Cambridge University Press.
- Heaton, J. (1990). *Classroom Testing*. Longman.
- Hopkins, K. D., Stanley, J. C., & Hopkins, B. R. (1990). Educational and psychological measurement and evaluation. Needham Heights, MA: Allyn & Bacon.
- Hughes, A. (2003). *Testing for Language Teachers* (2nd Ed.). Cambridge University Press.
- Islam, M.S., Hasan M.K., Sultana, S., Karim, A., & Rahman, M. M. (2021). English language assessment in Bangladesh today: principles, practices, and problem. *Language Testing in Asia*, 11(1), 2-21.
- Leta, D. (1990). *Achievement, Washback and Proficiency in School Leaving Examination: A Case of Innovation in an Ethiopian Setting* [Unpublished doctoral dissertation]. University of Lancaster.
- Manjarres, N. B. (2005). Washback of the foreign language test of the state examinations in Colombia: A case study. *Arizona Working Papers in SLAT*, 12(2005), 1-19.
- McCombes, S. (2022). *Descriptive Research Definition, Types, Methods & Examples*. <https://www.scribbr.com/methodology/descriptive-research/>.
- McNamara, T. (2000). *Language Testing*. Oxford: Oxford University Press.
- Moore, D. S., Notz, W. I, & Flinger, M. A. (2013). *The basic practice of statistics* (6th ed.). W. H. Freeman and Company.
- Nahdia, K., & Trisanti, N., (2019). Examining Washback of National Examination: EFL Teachers and Students' Perceptions. *Advances in Social Science, Education and Humanities Research*, 188, 273-277.
- Negede, D. (2002). *An Investigation of the Washback Effect of the Ethiopian School Leaving Certificate (ESLEC) English Language Examination* [Master's thesis, Addis Ababa University]. AAU Institutional Repository.

<http://etd.aau.edu.et/handle/123456789/28144>

- Oğuz, M. K., Demirogları, G. & Demirogları, B. (2020). Washback effect of higher education institution exam foreign language test on university candidates' attitudes. *African Educational Research Journal*, 8 (2), 121-130.
- Owusu, S. (2021). Washback Effect of High-Stakes English Language Tests on the Learning Behaviours of Ghanaian ESL Learners. *Journal of Education and Practice*, 12(30), 79-101.
- Pallant, J. (2020). *SPSS survival manual - a step by step guide to data analysis using SPSS for windows (version 10)* (7th ed.). Buckingham Open University Press. <https://doi.org/10.4324/9781003117452>.
- Pan, Y. (2014). Learner Washback Variability in Standardized Exit Tests. *The Electronic Journal for English as a Second Language: ESL-EJ*, 18(2), 1-30.
- Park, G-P, & Lee, H-W. (2006). The characteristics of effective English teachers as perceived by high school teachers and students in Korea. *Asian Pacific Education Review*, 7(2), 236-248.
- Punch, K.F. (2005). *Introduction to social research: quantitative and qualitative approaches* (2nd ed.). Sage.
- Rahman, Kh. A., Seraj, P.M.I, Hasan, Md. K., Namaziandost, E. & Tilwani, S.A. (2021). Washback of assessment on English teaching-learning practice at secondary schools. *Language Testing in Asia*, 11(12) 1-23.
- Rea-Dickins, P., & Scott, C. (2007). Washback from language tests on teaching, learning and policy: Evidence from diverse settings. *Assessment in Education: Principles, Policy & Practice*, 14(1), 1-7.
- Resnick, L. B. & Schantz, F. (2017). Testing, teaching, learning: who is in charge? *Assessment in Education: Principles, Policy & Practice*, 24(3), 424-432.
- Saif Sharzad (2006). Aiming for positive washback: a case study of international teaching assistants. *Language Testing*, 23 (1), 1-34.
- Salehi, H., & Yunus, Md. (2012). The washback effect of the Iranian universities entrance exam: teachers' insights. *Journal of Language Studies* 12(2), 609-628.
- Shih, C.-M. (2006). *Perceptions of the General English Proficiency Test and its washback: A case study at the two Taiwan technological institutes* [Doctoral dissertation, University of Toronto]. Tspace Repository. <https://hdl.handle.net/1807/119345>.
- Spratt, M. (2005). Washback and the classroom: the implications for teaching and learning of studies of washback from exams. *Language Teaching Research*, 9(1), 5-29.
- Stoneman, B. W. H. (2006). The impact of an exit English test on Hong Kong undergraduates: A study investigating the effects of test status on students' test preparation behaviours [Doctoral dissertation, Hong Kong Polytechnic University]. Polyu Electronic Theses Repository. <https://theses.lib.polyu.edu.hk/handle/200/5489>.
- Tayeb, Y., Aziz, M.S.A., Ismail, K. & Khan, A. (2014). The Washback Effect of the General Secondary English Examination (GSEE) on Teaching and Learning. <http://dx.doi.org/10.17576/GEMA-2014-1403-06>.

Taylor, L. (2005). Washback and Impact. *ELT Journal*, 59(2), 154-155.

Tokunaga, H. (2018). *Fundamental Statistics for the Social and Behavioral Sciences* (2nd ed.). Sage.

Wichadee, S. (2010). Characteristics of Effective English Language Teachers: The perspectives of Bangkok University Students.

http://www.bu.ac.th/knowledgecenter/epaper/jan_june2010/pdf/Page_01.pdf.

Zurrahmi, U., & Triastuti, A. (2022). Indonesian EFL students' perceptions of effective non-native English teachers. *Studies in English Language and Education*, 9(1), 299-317.

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).