A Comparative Study of English Performance in Online versus Traditional Education among Middle School Students in Zoucheng, China

Shuran Zhang[1], Norlizah Che Hassan[2]*, Roshafiza Binti Hassan[3], Xiao Huang[4], Yujie Long[5], Cuiyun Wang[6]

A R T I C L E   I N F O   **A B S T R A C T**

Aim: The primary objectives of this study were to assess the impact of two distinct educational models online education and traditional education on English performance skills scores, and to investigate whether gender has a significant influence on these outcomes. **Method:** The study sample comprised students from classes taught by the same instructor before and after the shutdown period. To evaluate the differences between the two educational models, an Independent-sample t-test employed, while gender-related calculations performed using MANOVA. **Results:** The findings revealed that English performance in online education (Mean = 65.73, p = .022) was superior to that in traditional education (Mean = 58.05). Specifically, scores for listening (Mean = 14.81, p = .050), reading (Mean = 28.61, p < .05), and vocabulary (Mean = 4.95, p = .029) were higher in the online education model compared to traditional education (Mean listening score = 13.63, Mean reading score = 23.27, Mean vocabulary score = 9.72). However, there was no significant difference in writing scores (p = .075) and grammar scores (p = .263) between the two models. Regarding gender, neither gender alone nor the interaction between gender and the education model significantly affected English performance (p = .077). Nonetheless, there were weakly significant differences observed in listening scores (p = .021) and writing scores (p = .033). **Conclusion:** This study highlights the relative advantages of online education and explores the impact of gender interaction with educational models on English performance. It also finds no significant statistical impact of gender differences on English performance, thereby contributing to the ongoing development of online English education.

[1] PhD Candidate, Faculty of Educational Studies, University Putra Malaysia, Selangor, Malaysia, 43400. ORCID iD: https://orcid.org/0009-0009-3911-7584, Email: gs65762@student.upm.edu.my

[2] Dr., Faculty of Educational Studies, University Putra Malaysia, Selangor, Malaysia, 43400. Email: norlizah@upm.edu.my

[3] Professor, Dr., Faculty of Educational Studies, University Putra Malaysia, Selangor, Malaysia, 43400. Email: roshafiza@upm.edu.my

[4] Zoucheng No. 12 Middle School, Zoucheng, China, 273500. Email: huangxiaodawn1992@163.com

[5] PhD Candidate, Faculty of Education, Universiti Kebangsaan Malaysia, Selangor, Malaysia, 43600. Email: P120115@siswa.ukm.edu.my

[6] PhD Candidate, Faculty of Educational Studies, University Putra Malaysia, Selangor, Malaysia, 43400. Email: gs64898@student.upm.edu.my

*Correspondence: norlizah@upm.edu.my

Shuran Zhang - Norlizah Che Hassan - Roshafiza Binti Hassan - Xiao Huang - Yujie Long - Cuiyun Wang / Eurasian Journal of Educational Research 111 (2024) 308-324

309

Introduction

The COVID-19 pandemic, which began in the spring of 2020, had a profound impact across all areas of life. Airborne transmission of SARS-CoV-2 has identified as a primary mode of spreading COVID-19 (Morawska & Cao, 2020). Evidence also suggests that the virus transmitted through simple conversation between infected and uninfected individuals (Anfinrud et al., 2020). Uninfected individuals are at increased risk of contracting the virus when they are in close proximity to infected persons within confined spaces (Dai & Zhao, 2020). Traditional education (TE), which typically involves fixed offline locations, staff, and teaching materials, inherently involves students being in a stable, often densely populated environment. Even though university students may change classrooms, the risk of infection remains high in such environments if an infected individual is present.

In response, to pandemic to mitigate the risk of infection, many countries suspended traditional classroom teaching. This shift significantly affected students' learning experiences. Effective study time is a well-established factor influencing academic performance, with a positive correlation between increased study time and better academic outcomes (Huebener, Kuger, & Marcus, 2017). To address these challenges, governments worldwide acted swiftly. On January 29, 2020, the Chinese government introduced the "stopping classes without stopping school" programme during the school shutdown (MEPRC, 2020). According to this initiative, schools were required to implement Online Education (OE) whenever infections detected in their area, ensuring the safety of students and staff. This measure remained in place until December 5, 2022, when the Chinese government reduced the COVID-19 risk level and lifted all control measures, allowing for a gradual return to traditional education.

The development of OE linked to the extent and advancement of information and communications technology. Initially, OE served as an ancillary tool for providing additional information to students. However, during the COVID-19 outbreak, OE mandated to replace TE and became the primary mode of instruction. This transition posed substantial challenges for both teachers and students, as the learning environment shifted abruptly to an online format. Despite these changes, the overarching educational objectives at national, school, and teacher levels remained largely unchanged, regardless of whether the education delivered online or traditionally. Nevertheless, OE, despite its notable flexibility, has faced criticism. Some scholars argue that OE is merely a temporary solution or even a failure (Wu, 2020). The effectiveness of OE compared to TE remains a subject of debate. Different subjects require varied teaching methods and educational goals.

English, as mandated by the Chinese Ministry of Education, is a compulsory subject from elementary school through to the second year of university. Some scholars argue that the primary goal of learning a second language is intercultural communication (Cui, 2020). However, there is a discrepancy between this goal and the reality for students below university level, where English education is often more exam-oriented, focusing on language rules such as phonetics, vocabulary, and grammar (Bi, 1998; Chen, 2008). Samovar and Porter (1991) suggest that communicative rules, which guide effective interpersonal interactions, should be emphasised, though they overshadowed by a focus on vocabulary and grammar in middle and high school due to the nature of English exams in China. Additionally, gender stereotypes in language learning persist, with males often

Shuran Zhang - Norlizah Che Hassan - Roshafiza Binti Hassan - Xiao Huang - Yujie Long - Cuiyun Wang / Eurasian Journal of Educational Research 111 (2024) 308-324

310

perceived as stronger in mathematics and females in languages like Chinese and English (Wang, 2021). This stereotype can place undue pressure on students.

Thus, further investigation into the impact of gender on academic performance warranted. This study aims to evaluate the influence of the two educational models and the role of gender in these contexts. We compared two groups of students who taught by the same teacher and assessed using the same English test paper. Group A completed their studies and test during the COVID-19 outbreak, while Group B did so under traditional education conditions. The study has two primary objectives: 1) to identify the differences in English performance, including overall performance and scores in each English skill area, between the two groups; and 2) to examine whether there are significant gender differences in English performance (overall and for each English skill) between the two groups.

Literature Review

Online Education (OE), as a relatively new educational model, offers the significant advantage of overcoming limitations of time and space that are inherent in Traditional Education (TE) (Means et al., 2009). In this context, OE refers to courses delivered over the Internet during the school closures due to COVID-19. While the organisational aspects of classes, learning content, and schedules remained consistent, OE differed from TE in that it relied entirely on digital networks and devices, with the entire teaching process dependent on online platforms. Modern education increasingly emphasises a blend of self-directed and collaborative learning, marking a shift from teacher-led to student-led instruction. OE has substantially facilitated this shift. However, compared to TE, OE generally offers less interaction with teachers and peers due to changes in educational space and tools, even though lecture content has remained relatively unchanged. The learning process in OE places greater reliance on students' self-motivation, parental involvement, and other factors. Thus, while OE enhances interactive learning, TE traditionally emphasises interaction with others (Anggrawan & Jihadil, 2018). The effectiveness of these two educational models remains a topic of considerable debate. Academic performance frequently used as a benchmark for evaluating educational effectiveness (Bintoro, Lestari, & Aini, 2022; Longhu, 2022; Wieser & Seeler, 2018).

Research typically employs experimental or quasi-experimental methods to explore the relationship between educational models and academic outcomes. Although OE has been shown to be effective (Shachar & Neumann, 2003), some studies suggest that it may not be as effective as TE. For instance, Coates et al. (2004) found that students using TE performed better in understanding college economics compared to those learning online, even after accounting for the endogenous factor of "student self-selection." Furthermore, OE often results in more dispersed and polarised academic achievement (Shen et al., 2007), potentially affecting future performance (Bettinger et al., 2017). Similar trends have observed in younger students; for example, Sipahutar (2023) reported that high school students performed better in TE. Eighth graders with a solid math foundation scored lower in online courses compared to their peers with weaker math backgrounds in traditional classes (Ediwarman, 2023; Heppen et al., 2017). One contributing factor is the lack of teacher feedback and tailored instruction in OE (Eom, Wen, & Ashill, 2006). Improvements in teacher attendance could significantly enhance the effectiveness of OE (Zhao et al., 2005).

Unfortunately, the interaction between students and teachers in TE cannot fully replicated in OE due to technological and pedagogical limitations. Conversely, some studies report no significant difference in effectiveness between OE and TE. McPhee and Söderström (2012) found that neither educational model, students' nationality, nor study time significantly affected academic performance. Similar findings reported in STEM fields, with no significant differences observed between the two models (Paul & Jefferson, 2019; Wang et al., 2019). Dendir (2019) reached similar conclusions after considering the self-selection bias. However, these studies often involve small sample sizes (e.g., McPhee & Söderström's 167 participants and Wang et al.'s 64 participants), which may not be representative of broader population characteristics. Variations in study populations and non-probability sampling could contribute to these inconclusive results. On the other hand, some researchers have found that OE students perform better in TE (Alam & Parvin, 2021; Clark et al., 2021; Dolfsma, van der Eijk, & Mahdi, 2022; Meeter, 2021; Spitzer & Musslick, 2021). Particularly in schools with high-quality students and fewer vulnerable individuals, online learners showed greater progress (Aguaded et al., 2023; Meeter, 2021).

However, this advantage appears confine to academic performance, with aptitude and practicum test scores showing less improvement (Alam & Parvin, 2021; Mukongo, Otafudu, & Toko, 2023). Early predictions suggested OE might lead to "learning loss" during the pandemic (Brown et al., 2020), but this has been contradicted in some cases. The short duration of lockdowns in certain countries, which limits observation time and contextual differences such as technological development, cultural factors, and family backgrounds, must consider when interpreting these findings (Gore et al., 2021). In terms of English education, studies have similarly varied. English performance and language skills (listening, speaking, reading, and writing) show mixed results regarding the effectiveness of OE versus TE. Some research indicates that, OE perceived effective learning method by students (Zakarneh, 2018). However, Xu and Jaggars (2011) found that students performed poorly in introductory-level English online courses, attributing this to limitations in online teamwork and communication.

Additionally, some students found traditional classes more conducive to focusing on teacher instruction (Luo, 2024; Rahmat, 2022). OE has reported positively affect specific English skills such as listening (Amir & Kang, 2018; Jabber, Charfeddine, & Abbas, 2023; Salsabila Gita & Utami, 2021). The increased availability of online materials allows for more reading opportunities, which can enhance motivation (Jabeen & Thomas, 2015; Khodaparast & Ghafournia, 2015). Writing, being the most complex skill, often relies on teacher feedback, which in OE can be slower, though generally, there is satisfaction with online writing instruction (Sheerah et al., 2022). The effectiveness of OE for speaking skills remains contentious, with some arguing that reduced teacher feedback in OE hampers speaking proficiency (Jabeen & Thomas, 2015; Vargas-Quesada et al., 2023). Research on the relationship between gender and English performance (EP) has produced inconsistent results. Some studies indicate a weak relationship between gender and EP (Danuwijaya & Roebianto, 2020; Ossai, 2019), while others suggest a stronger correlation (Adams & Simmons, 2019; Schleeter et al., 2019). Few studies have explored how gender affects EP across different educational models. Some research suggests no significant gender differences in academic performance when not restricting the analysis to specific courses. Similarly, some studies report that while male students show similar performance across OE and TE, female students perform worse in OE compared to TE (Flanagan, 2012).

Research Methodology Design

Participants

The research sample comprised 9th-grade students from the 2022 and 2023 cohorts at a middle school in Zoucheng, China. Total 213 students participated; the same English teacher taught all of whom. The 2022 cohort, referred to as Group I, included 109 students across two classes, while the 2023 cohort, known as Group II, comprised 104 students from two different classes. To assess the difference in English performance between Online Education (OE) and Traditional Education (TE), the English midterm exam results used as the primary measure. Group I, who studied and took the exam online during the lockdown period, represents the OE group. In contrast, Group II, who studied and completed their exam on campus, represents the TE group. Table 1 details the gender distribution among the 213 students: there were 108 females and 105 males. Specifically, Group I consisted of 56 males and 53 females, while Group II included 49 males and 55 females. Due to the relatively small size of the population, random sampling has not employed in this study, meaning the population size was equivalent to the sample size.

Table 1

Sociodemographic Characteristics of Participants

		Group		Total
		I	**II**	
Gender	Male	49	56	105
	Female	55	53	108
Total		104	109	213

Data Collection

This study employed a quasi-experimental research approach to measure the effectiveness of Online Education (OE) versus Traditional Education (TE) by evaluating English performance (EP). Both groups of students instructed by the same teacher and used nearly identical teaching materials and the same exam paper. The primary distinction between the groups lay in the educational models they experienced, thereby minimising variations related to teaching and assessment. The study assessed five components of English skills: listening, reading, writing, grammar, and vocabulary. These components collectively referred to as English skills grades. The listening section included multiple-choice questions (Q1 - Q15) and fill-in-the-blank questions (Q31 - Q35). The reading section comprised multiple-choice questions (Q16 - Q30) and task-based reading questions (Q61 - Q65). Grammar was tested through fill-in-the-blank (Q44 - Q53) and completion tests (Q54 - Q60). The vocabulary section consisted of questions (Q36 - Q43) involving filling in blanks with the initial letter of words and translating Chinese words into English. The writing skill assessed through a single question requiring a written response on a fixed topic. The exam paper scored out of 100 points, with the following distribution: 20 marks for listening, 40 for reading, 17 for grammar, 8 for vocabulary, and 15 for writing.

Shuran Zhang - Norlizah Che Hassan - Roshafiza Binti Hassan - Xiao Huang - Yujie Long - Cuiyun Wang / Eurasian Journal of Educational Research 111 (2024) 308-324

313

Methods

To evaluate the differences in effectiveness between OE and TE, Descriptive Statistics and Independent-sample t-tests employed. The observation values for both groups were independent and met the Homogeneity of Variance Test at a 0.05 level of significance. Although the data for EP (p < .05) and individual English skills scores (p < .05) did not pass the Shapiro-Wilk test for normality, the kurtosis and skewness values were within acceptable limits (absolute kurtosis < 10 and absolute skewness < 3). This, combined with the histogram and Q-Q plot, suggested that the data approximately followed a normal distribution, thus meeting the criteria for the t-test (Reiss, Thomas, & Reiss, 1997). All statistical analyses performed using IBM SPSS Statistics 27. Additionally, Univariate Analysis of Variance (ANOVA) and Multivariate Analysis of Variance (MANOVA) were utilised to examine the impact of OE and TE on English skills scores and EP across genders. These analyses determined whether there was an interaction between the educational models and gender. Box's M test indicated that the data met the criteria for Homogeneity of variance-covariance (p = .068), and Levene's test confirmed equal variances among the groups of independent variables (p > .05).

Findings

English Performance between Online and Traditional Education

Table 2 presents the results of the descriptive statistics and Independent-sample t-test. Descriptive statistics included the mean scores and standard deviations for both groups. The mean scores reveal the central tendencies, while standard deviation (SD) measures dispersion. The mean score for Group I (Mean = 65.73, N = 108, SD = 24.41) was higher than that for Group II (Mean = 58.05, N = 104, SD = 24.20). According to China's performance classification standards, only Group I achieved a mean score above the passing mark of 60. Despite similar SDs for both groups, with a difference of approximately 0.21, the score distributions varied. Most students in Group I scored between 43.31 and 90.14, whereas Group II students' scores ranged from 33.85 to 82.25. Scores above 90 considered excellent, and those between 80 and 90 deemed good. Thus, the highest scores in Group I were at a higher level compared to Group II, indicating a significant difference between OE and TE. The Independent-sample t-test results (Table 2) showed a mean difference in EP between the two groups with t = 2.30, p = .022. The effect size (đ) was 0.32, reflecting a small effect size (Cohen, 1992). These results support the conclusion that there is a significant difference in EP between the OE and TE groups at the 0.05 level of significance.

Table 2

Results of Independent-Sample t-Test on EP by Groups

Groups	n	Mean	SD	t	p	$\hat{d}$
				2.30	.022	.32
I	109	65.73	24.41			
II	104	58.05	24.20			

Notes: *n = Sample Size, SD = Standard Deviation, $\hat{d}$ = Effect Size.*

Table 3 presents the results of the descriptive statistics and Independent-sample t-tests for English skills scores across the Online Education (OE) and Traditional Education (TE) groups. The data indicate that the mean grammar score for the OE group (Mean = 8.90, N = 109, SD = 5.28) was lower than that of the TE group (Mean = 9.72, N = 104, SD = 5.42). Conversely, the OE group outperformed the TE group in the other four English skills categories. Although the difference in the mean vocabulary score between OE (Mean = 4.95, N = 109, SD = 2.91) and TE (Mean = 4.03, N = 104, SD = 3.17) was only 0.92, it is noteworthy given that the vocabulary test was scored out of 08 marks. This difference represents a significant percentage of the total vocabulary marks, suggesting a meaningful disparity between the two groups. Independent-sample t-tests revealed significant mean differences in listening (t = 1.97, p = .050), reading (t = 4.23, p < .05), and vocabulary scores (t = 2.20, p = .029) between the two groups. The effect sizes for these differences were small for listening ($\hat{d}$ = .27) and vocabulary scores ($\hat{d}$ = .30), while the effect size for the reading score was medium ($\hat{d}$ = .58). No significant differences were found in grammar (t = -1.12, p = .263) and writing scores (t = 1.79, p = .075) between the groups at the 0.05 level of significance.

Table 3

Results of Independent-Sample t-Test on English Skill Scores by Groups

	Groups	n	Mean	SD	t	p	$\hat{d}$
Listening					1.97	.050	.27
	I	109	14.81	4.17			
	II	104	13.63	4.57			
Reading					4.23	.000	.58
	I	109	28.61	9.83			
	II	104	23.27	8.511			
Grammar					-1.12	.263	-.15
	I	109	8.90	5.28			
	II	104	9.72	5.42			
Vocabulary					2.20	.029	.30
	I	109	4.95	2.91			
	II	104	4.03	3.1668			
Writing					1.79	.075	.25
	I	109	8.46	4.0950			
	II	104	7.40	4.4928			

Notes: *n = Sample Size, SD = Standard Deviation, $\hat{d}$ = Effect Size.*

Gender Differences in English Performance between Online and Traditional Education

Table 4 showed the mean and standard deviation of students' EP of different gender in different groups. As can be seen, male students' mean of EP (Mean = 67.1, SD = 3.22) was higher than female students' (Mean = 64.26, SD = 3.42) in Group I, but surprisingly, Group II was the exact opposite. In order to explore whether gender actually made the differences, we used Univariate Analysis of Variance to analyse the effects of gender, education, interaction (Gender*Group) on EP. According to Table 5, there was no significant effect on EP by interaction (Gender*Group), F = 3.16, p = .077, η^2 = .015. Meanwhile, gender's main effect was not significant, F=0.85, p = .358, η^2 = .004. However,

Shuran Zhang - Norlizah Che Hassan - Roshafiza Binti Hassan - Xiao Huang - Yujie Long - Cuiyun Wang / Eurasian Journal of Educational Research 111 (2024) 308-324

315

in Table 5, the main effect of Group was statistically significant, F = 5.66, p = .018, η^2 = .026, which further validated the conclusions of 4.1 in this section.

Table 4

Results of Descriptive Statistics of EP

	Group I			Group II		
Gender	n	Mean	SD	n	Mean	SD
Male	56	67.11	3.22	49	53.31	3.49
Female	53	64.26	3.42	55	62.27	3.16

Notes: *n = Sample Size, SD = Standard Deviation, $\hat{d}$ = Effect Size.*

Table 5

Results of Between-Subject Effects (at p<.05 level)

Factor	F	p	η^2
Gender	.85	.358	.004
Group	5.66	.018	.026
Interaction (Gender*Group)	3.16	.077	.015

R Squared = .043 (Adjust R Squared = .029)

Table 6 presents the results of the Multivariate Analysis of Variance (MANOVA), revealing a significant interaction effect of gender and education model on English skills scores (F = 2.40, p = .039, η^2 = .055). This interaction accounted for 5.5% of the variance in the overall English skills scores, indicating a small effect size. As detailed in Table 7, this interaction effect was significant for the listening score (p = .025) and the writing score (p = .021). Specifically, 2.5% of the variance in listening scores and 2.1% of the variance in writing scores attributed to the interaction between gender and education model. Conversely, there were no statistically significant interaction effects for reading (p = .307), grammar (p = .107), and vocabulary scores (p = .054).

Table 6

Significant Multivariate Effects (at p<.05 level)

Factor	Wilk's Lambda	F	p	η^2
Gender	.934	2.91	.015	.066
Group	.662	20.91	.000	.338
Interaction (Gender*Group)	.945	2.40	.039	.055

Table 7

*Significant Between-Subject Effects for Gender*Group (at p<.05 level)*

Dependent Variable	df	F	Sig.	η^2
Listening	1	5.38	.021	.025
Reading	1	1.05	.307	.005
Grammar	1	2.62	.107	.012
Vocabulary	1	3.76	.054	.018
Writing	1	4.59	.033	.021

R Square = .046 (Adjusted R Square = .033)

To delve deeper into the results, a Univariate Analysis of Variance (ANOVA) conducted to examine the simple effects. The analysis, presented in Table 8, indicated a significant effect of the education model on listening scores (F = 4.07, p = .045). This suggests that the education model had a differential impact on listening scores. However, there was no significant effect of gender on listening scores (F = 0.01, p = .941). Further, both the main effect of the education model and the interaction effect between gender and education model (F = 5.38, p = .021) were significant. Table 9 provides additional insight, revealing that female students' listening scores did not differ significantly between OE and TE (p = .058). In contrast, male students exhibited a significant change in their listening scores between the two education models (p = .003), with OE leading to higher scores compared to TE (mean difference = 2.58).

Table 8

Results of Between-Subject Effects on Listening Scores (at p<.05 level)

Factor	F	p	η^2
Gender	.01	.941	.000
Group	4.07	.045	.019
Interaction (Gender*Group)	5.38	.021	.025
R Squared = .043 (Adjust R Squared = .029)			

Table 9

Results of Simple Effects Analysis in Listening Score (Gender&Group)

Gender	Factor I - Factor J	MD	SE	P
Male	I-II	2.58*	.85	.003
Female	I-II	-.18	.84	.830

Notes: *MD = Mean Difference, SE = Standard Error.*

Table 10 indicates that neither the education model (F = 3.65, p = .058) nor gender (F = 2.42, p = .122) had a significant effect on writing scores. This suggests that neither the type of education model nor gender alone produced a differential effect on writing scores. Given the significant interaction effect (Gender*Group), further simple effects analysis was conducted, as detailed in Tables 11 and 12. The results showed that female students did not experience a significant interaction effect between OE and TE on writing scores (p = .869). In contrast, a significant interaction effect was observed for male students (p = .005), who achieved higher writing scores in OE compared to TE (mean difference = 2.36). Additionally, within the OE model, there were no significant differences between male and female students in writing scores (p = .675). However, in the TE model, male students scored lower than female students did on the writing test.

Table 10

Results of Between-Subject Effects on Writing Score (at p<.05 level)

Factor	F	p	η^2
Gender	2.42	.122	.011
Group	3.65	.058	.017
Interaction (Gender*Group)	4.59	.033	.021
R Squared = .046 (Adjust R Squared = .033)			

Table 11

Results of Simple Effects Analysis in Writing Score (Gender Group)

Gender	Group	MD	SE	P
Male	I-II	2.36*	.83	.005
Female	I-II	-.14	.82	.869

Notes: *MD = Mean Difference, SE = Standard Error.*

Table 12

Results of Simple Effects Analysis in Writing Score (Gender Group)

Gender	Group	MD	SE	P
Group I	Male-Female	.34	.81	.675
Group II	Male-Female	-2.15*	.83	.011

Notes: *MD = Mean Difference, SE = Standard Error.*

Although the overall main effect of gender on English Skills Scores was statistically significant (F = 2.91, p = .015, η² = 0.07), the simple effects analysis presented in Table 13 revealed that gender did not show a statistically significant impact on any of the individual English Skills Scores after controlling for the education model (p < .05). This suggests that gender may not consistently influence each dependent variable in a uniform direction. Conversely, the main effect of the education model on English Skills Scores was statistically significant (F = 20.91, p = .000, η² = 0.34). Table 14 further illustrates that, after controlling for gender, significant differences were found between OE and TE in the listening score (F = 4.07, p = .045, η² = 0.02), reading score (F = 18.46, p = .000, η² = 0.08), and vocabulary score (F = 5.42, p = .021, η² = 0.03). These results corroborate the previous findings in this study, confirming the differential impact of the education model on these specific English skills.

Table 13

Simple Effect for Gender (at p<.05 level)

Dependent Variable	df	F	Sig.	η^2
Listening	1	.01	.941	.000
Reading	1	1.33	.251	.006
Grammar	1	.001	.975	.000
Vocabulary	1	3.07	.081	.014
Writing	1	2.42	.122	.011

R Square = .046 (Adjusted R Square = .033)

Table 14

Simple Effect for Group (at p<.05 level)

Dependent Variable	df	F	Sig.	η^2
Listening	1	4.07	.045	.019
Reading	1	18.46	.000	.081
Grammar	1	1.20	.274	.006
Vocabulary	1	5.42	.021	.025
Writing	1	3.65	.058	.017

R Square = .046 (Adjusted R Square = .033)

Shuran Zhang - Norlizah Che Hassan - Roshafiza Binti Hassan - Xiao Huang - Yujie Long - Cuiyun Wang / Eurasian Journal of Educational Research 111 (2024) 308-324

318

Discussion

In this study, we evaluated the effects of different educational models OE and TE on the English performance and English skills scores of 9th-grade middle school students. We also examined how these effects varied by gender. Our findings indicated significant differences between students participating in online versus face-to-face (F2F) education in terms of EP, reading scores, listening scores, and vocabulary scores, with online students generally achieving higher scores compared to their traditional class counterparts. Writing and grammar scores, however, remained comparable across both education models. While gender and the interaction between gender and education model (Gender*Group) did not affect EP, they had a small effect size on English skills scores, particularly in listening and writing. Despite limitations such as a small sample size and insufficiently varied test questions, especially for vocabulary and grammar, these findings are consistent with some existing research in this area. The results suggested that OE is sufficiently effective and integrated into regular school education even after the mandatory online education period ends. Many studies in this field have predominantly utilised Likert scales to assess perceptions of teachers and students, with few comparing statistical differences in performance, particularly in English. MANOVA commonly used to compare gender differences, less frequently applied to compare academic performance between OE and TE, making this study innovative in its analytical approach.

In this research, EP calculated as the sum of individual English skill scores. Since three out of the five English skill scores improved with OE and the remaining two stayed essentially the same, it was inevitable that EP showed an overall improvement. The enhancements in listening, reading, and vocabulary scores associated with OE attributed to the pedagogical advantages of these skills. Second language learners are encouraged to view vocabulary as an integral part of reading rather than as isolated elements (Wang, 1998). Jabeen and Thomas (2015) noted that rich reading materials could stimulate students' interest in reading and boost their motivation, thereby enhancing the effectiveness of online learning. Traditional education limits students to paper-based materials within the classroom, and in China, students prohibited from bringing mobile phones and laptops to school, with schools unable to provide electronic devices for each student. Consequently, reading materials are often restricted. OE, by providing a greater variety of materials suited to students' abilities, facilitates vocabulary development through increased reading opportunities. This bidirectional benefit enhances reading scores, as a larger vocabulary positively affects reading competence (Duan, Jiang, & Mei, 2000). Similarly, language characteristics, individual differences, and the learning environment influence listening skills. Factors such as pronunciation and speaking speed can challenge students' learning. Online education offers more frequent practice opportunities and a wider range of materials, which can help overcome these challenges.

Additionally, the improvement in scores linked to increased parental involvement during OE. Research has shown a positive correlation between parental involvement and academic performance (Bai et al., 2021; Pajarianto, 2020). Emotional support and non-punitive physical participation from parents can aid in the successful completion of assignments. For instance, memorising English words, which is often challenging for students, approached through rote learning in China (Wang, 1998). However, procrastination and laziness among some middle school students can delay this memorisation. In the 9th grade, where there is significant

Shuran Zhang - Norlizah Che Hassan - Roshafiza Binti Hassan - Xiao Huang - Yujie Long - Cuiyun Wang / Eurasian Journal of Educational Research 111 (2024) 308-324

319

pressure to advance to higher education, students might use activities to alleviate stress (Wang, 2008). Family emotional support and supervision can mitigate these issues. During the TE period post-shutdown, work and social interactions affected parental involvement in education. However, during the shutdown, government-imposed stay-at-home orders allowed many parents the opportunity to participate in online education. In Shandong Province, known for its strong educational values, nearly 90% of students accompanied by their parents during this time (Bai et al., 2021). Parents relayed information from schools and took on some teaching responsibilities, which might explain why EP in OE was higher than in TE.

Conclusion and Limitations

In conclusion, both the results of this study and previous research suggest that OE has potential as a component of school education. While OE has its drawbacks, these should not lead to its complete rejection. Embracing the strengths of different educational models to address deficiencies is crucial for the advancement of education. The unique characteristics of OE during shutdown time, such as increased parental involvement, should be emphasised to enhance students' academic performance. Compared with the gender factor, the class and relatively low socio-economic status become non-negligible to learning loss of OE (Aguaded et al., 2023). English, being a language, should be taught in a manner consistent with its inherent characteristics rather than through rote memorisation. The study found no significant gender differences in EP, nor did the interaction between gender and education model affect EP. However, there was a notable interaction effect on English skill scores, although the effect size was small, accounting for only 5.5% of the variance. This interaction was statistically significant only for listening and writing scores, not for reading, grammar, or vocabulary. Further analysis revealed that male students showed significant improvements in listening and writing scores in OE compared to TE. This may be due to physiological and psychological differences between genders, such as men's lesser ability to recognise negative emotions (Li, Yuan, & Lin, 2008) and their greater adaptability to sudden changes (Denies et al., 2022). Alternatively, this attributed to the small sample size or other uncontrolled factors. The study's limitations include a restricted sample size, limited test content, variability in student quality, and disparities in students' information technology capabilities. The focus on classes taught by the same teacher before and after the COVID-19 shutdown intended to reduce variations in teaching and test content but resulted in a limited sample size. Additionally, differences in student quality between the groups could affect results. The test papers had a small number of questions, particularly for vocabulary. Technological advancements may also reveal gaps in students' IT skills for online learning. Future research should address these limitations, investigate other relevant factors such as family involvement, and explore hybrid-teaching methods that combine OE and TE.

References

Adams, A.-M., & Simmons, F. R. (2019). Exploring individual and gender differences in early writing performance. *Reading and Writing, 32*(2), 235-263. https://doi.org/10.1007/s11145-018-9859-0

Aguaded I., Vizcaíno-Verdú A., García-Prieto V., & de-Casas-Moreno P. (2023). The Impact of Post-Pandemic Learning Loss on Education Development: A Systematic Review.

Shuran Zhang - Norlizah Che Hassan - Roshafiza Binti Hassan - Xiao Huang - Yujie Long - Cuiyun Wang / Eurasian Journal of Educational Research 111 (2024) 308-324

320

Review of Communication Research, 11, 172–189. https://doi.org/10.5680/RCR.V11.7

Alam, G. M., & Parvin, M. (2021). Can online higher education be an active agent for change? —comparison of academic success and job-readiness before and during COVID-19. *Technological Forecasting and Social Change, 172*, 121008. https://doi.org/10.1016/j.techfore.2021.121008

Amir, S., & Kang, M. A. (2018). Research in Use of Information & Communication Technologies (ICT) for Developing Listening Comprehension Competency in Foreign/second Languages: A Review of Selected Tools. *International Journal of Social Sciences & Humanities (IJSSH), 3*(1), 44-53. http://ijssh.ielas.org/index.php/ijssh/article/view/24

Anfinrud, P., Stadnytskyi, V., Bax Christina, E., & Bax, A. (2020). Visualizing Speech-Generated Oral Fluid Droplets with Laser Light Scattering. *New England Journal of Medicine, 382*(21), 2061-2063. https://doi.org/10.1056/NEJMc2007800

Anggrawan, A., & Jihadil, Q. S. (2018). Comparative analysis of online e-learning and face to face learning: an experimental study. In *2018 Third International Conference on Informatics and Computing (ICIC)* (pp. 1-4). IEEE. https://doi.org/10.1109/IAC.2018.8780495

Bai, R., Zhang, W. Y., Guan, Y., & Xie, Q. S. (2021). An Analysis of the Transformation Path of Home-school Collaboration in the Post-epidemic Era——based on a Questionnaire Survey of 1,440 Families During the Period of "suspension of Classes but Not Learning". *China Educational Technology, 410*(03), 30-37. http://ir.library.ouchn.edu.cn/handle/39V7QQFX/49022

Bettinger, E. P., Fox, L., Loeb, S., & Taylor, E. S. (2017). Virtual Classrooms: How Online College Courses Affect Student Success. *American Economic Review, 107*(9), 2855–2875. https://doi.org/10.1257/aer.20151193

Bi, J. W. (1998). Intercultural Communication Studies and Second Language Teaching. *Language Teaching and Linguistic Studies,* (01), 10-24. https://cstj.cqvip.com/Qikan/Article/Detail?id=1002380369

Bintoro, T., Lestari, I., & Aini, I. N. (2022). Analyzing learners' needs and designing digital comic media to improve student learning outcomes. *Educational Sciences: Theory & Practice, 22*(1), 129-140. https://jestp.com/menuscript/index.php/estp/article/view/1563

Brown, N., Te Riele, K., Shelley, B., & Woodroffe, J. (2020). *Learning at home during COVID-19: Effects on vulnerable young Australians*. Hobart: University of Tasmania, Peter Underwood Centre for Educational Attainment. https://wscf.org.au/wp-content/uploads/2021/07/Learning-at-home-during-COVID-19-Effects-on-vulnerable-young-Australians.pdf

Chen, G. H. (2008). Analysis and Recommendations on the Current Situation and Policies of English Education in China. *Foreign Languages in China,* (02), 4-6+14.

Clark, A. E., Nong, H., Zhu, H., & Zhu, R. (2021). Compensating for academic loss: Online learning and student performance during the COVID-19 pandemic. *China Economic Review, 68*, 101629. https://doi.org/10.1016/j.chieco.2021.101629

Coates, D., Humphreys, B. R., Kane, J., & Vachris, M. A. (2004). "No significant distance" between face-to-face and online instruction: evidence from principles of economics. *Economics of Education Review, 23*(5), 533-546. https://doi.org/10.1016/j.econedurev.2004.02.002

Shuran Zhang - Norlizah Che Hassan - Roshafiza Binti Hassan - Xiao Huang - Yujie Long - Cuiyun Wang / Eurasian Journal of Educational Research 111 (2024) 308-324

321

Cohen, J. (1992). A Power Primer. *Psychological Bulletin, 112*(1), 155-159. https://doi.org/10.1037/0033-2909.112.1.155

Cui, Y. H. (2020). The Goal of Teaching Chinese as a Second Language: to Cultivate the Chinese Intercultural Communication Competence. *Language Teaching and Linguistic Studies, 4*, 25-36. https://www.linglab.cn/news/1556

Dai, H., & Zhao, B. (2020). Association of the infection probability of COVID-19 with ventilation rates in confined spaces. *Building Simulation, 13*(6), 1321-1327. https://doi.org/10.1007/s12273-020-0703-5

Danuwijaya, A. A., & Roebianto, A. (2020). Performance differences by gender in English reading test. *Jurnal Penelitian dan Evaluasi Pendidikan, 24*(2), 190-197. https://doi.org/10.21831/pep.v24i2.34344

Dendir, S. (2019). Performance differences between face-to-face and online students in economics. *Journal of Education for Business, 94*(3), 175-184. https://doi.org/10.1080/08832323.2018.1503586

Denies, K., Heyvaert, L., Dockx, J., & Janssen, R. (2022). Mapping and explaining the gender gap in students' second language proficiency across skills, countries and languages. *Learning and Instruction, 80*, 101618. https://doi.org/10.1016/j.learninstruc.2022.101618

Dolfsma, W., van der Eijk, R., & Mahdi, S. (2022). Front-loading the front-end of innovation. *The Journal of Modern Project Management, 10*(1), 140-149. https://journalmodernpm.com/manuscript/index.php/jmpm/article/view/JMPM02909

Duan, H. F., Jiang, Z. C., & Mei, W. G. (2000). An Initial Exploration of the Structure of Secondary School Students' English Reading Ability and Its Development. *Journal of Psychological Science,* (05), 547-551, 637. https://doi.org/10.16719/j.cnki.1671-6981.2000.05.009

Ediwarman, E. (2023). The Global teaching practices of Bahasa Indonesia. *Eurasian Journal of Applied Linguistics, 8*(3), 312-320. https://ejal.info/menuscript/index.php/ejal/article/view/436

Eom, S. B., Wen, H. J., & Ashill, N. (2006). The Determinants of Students' Perceived Learning Outcomes and Satisfaction in University Online Education: An Empirical Investigation. *Decision Sciences Journal of Innovative Education, 4*(2), 215-235. https://doi.org/10.1111/j.1540-4609.2006.00114.x

Flanagan, J. (2012). Online Versus Face-To-Face Instruction: Analysis of Gender and Course Format in Undergraduate Business Statistics Courses. *Academy of Business Research, 2*, 93-101. https://ssrn.com/abstract=2627101

Gore, J., Fray, L., Miller, A., Harris, J., & Taggart, W. (2021). The impact of COVID-19 on student learning in New South Wales primary schools: an empirical study. *The Australian Educational Researcher, 48*(4), 605-637. https://doi.org/10.1007/s13384-021-00436-w

Heppen, J. B., Sorensen, N., Allensworth, E., Walters, K., Rickles, J., Taylor, S. S., et al. (2017). The Struggle to Pass Algebra: Online vs. Face-to-Face Credit Recovery for At-Risk Urban Students. *Journal of Research on Educational Effectiveness, 10*(2), 272-296. https://doi.org/10.1080/19345747.2016.1168500

Huebener, M., Kuger, S., & Marcus, J. (2017). Increased instruction hours and the widening gap in student performance. *Labour Economics, 47*, 15-34. https://doi.org/10.1016/j.labeco.2017.04.007

Jabber, H. N., Charfeddine, B., & Abbas, H. J. (2023). Evaluation of FGF23, C-reactive Protein and Lipid Profile as Predictors for Cardiovascular Diseases in Patients

Shuran Zhang - Norlizah Che Hassan - Roshafiza Binti Hassan - Xiao Huang - Yujie Long - Cuiyun Wang / Eurasian Journal of Educational Research 111 (2024) 308-324

322

with Type 2 Diabetes. *The Review of Diabetic Studies, 19*(3), 145-157. https://diabeticstudies.org/menuscript/index.php/RDS/article/view/371

Jabeen, S. S., & Thomas, A. J. (2015). Effectiveness of Online Language Learning. In *Proceedings of the World Congress on Engineering and Computer Science* (Vol. 1, pp. 21-23). WCECS. https://www.iaeng.org/publication/WCECS2015/WCECS2015_pp297-301.pdf

Khodaparast, F., & Ghafournia, N. (2015). On the Effect of Online/Offline Approaches on Vocabulary Achievement. *International Journal of Education and Research, 3*(9), 269-280. https://www.ijern.com/journal/2015/September-2015/22.pdf

Li, H., Yuan, J., & Lin, C. (2008). The neural mechanism underlying the female advantage in identifying negative emotions: An event-related potential study. *NeuroImage, 40*(4), 1921-1929. https://doi.org/10.1016/j.neuroimage.2008.01.033

Longhu, C. (2022). The Discussion of Ziben zhuyi (capitalism) in China's Debate on Socialism (1920-1921). *Cultura: International Journal of Philosophy of Culture and Axiology, 19*(1), 75-86. https://culturajournal.com/submissions/index.php/ijpca/article/view/171

Luo, Y. (2024). The Time Value and Practice Path of the Integration of Sports Spirit into Ideological and Political Teaching in Colleges and Universities. *Revista de Psicología del Deporte (Journal of Sport Psychology), 33*(1), 287-294. http://rpd-online.com/index.php/rpd/article/view/1550

McPhee, I., & Söderström, T. (2012). Distance, online and campus higher education: reflections on learning outcomes. *Campus-Wide Information Systems, 29*(3), 144-155. https://doi.org/10.1108/10650741211243166

Means, B., Toyama, Y., Murphy, R., Bakia, M., & Jones, K. (2009). *Evaluation of Evidence-Based Practices in Online Learning: A Meta-Analysis and Review of Online Learning Studies*. US Department of Education. https://www2.ed.gov/rschstat/eval/tech/evidence-based-practices/finalreport.pdf

Meeter, M. (2021). Primary school mathematics during the COVID-19 pandemic: No evidence of learning gaps in adaptive practicing results. *Trends in Neuroscience and Education, 25*, 100163. https://doi.org/10.1016/j.tine.2021.100163

MEPRC. (2020). *Ministry of Education: Use of Online Platforms, "Stopping Classes Without Stopping School"*. Ministry of Education of the People's Republic of China. http://www.moe.gov.cn/jyb_xwfb/gzdt_gzdt/s5987/202001/t20200129_416993.html

Morawska, L., & Cao, J. (2020). Airborne transmission of SARS-CoV-2: The world should face the reality. *Environment International, 139*, 105730. https://doi.org/10.1016/j.envint.2020.105730

Mukongo, D. K., Otafudu, O. O., & Toko, W. B. (2023). On The Ultra-Quasi-Tight Extensions. *Math. Appl, 12*, 101-114. https://mathsapplication.com/wp-content/uploads/2023/07/10.13164.ma_.2023.07.pdf

Ossai, P. A. U. (2019). Gender Variation in Undergraduate Students' Achievement in English Language. *Journal of Science Technology & Education, 7*(3), 130-136.

Pajarianto, H. (2020). Study From Home in the Middle of the COVID-19 Pandemic: Analysis of Religiosity, Teacher, and Parents Support Against Academic Stress. *Talent Development & Excellence, 12*(2), 1791-1807. https://www.researchgate.net/publication/341805032

Paul, J., & Jefferson, F. (2019). A Comparative Analysis of Student Performance in an Online vs. Face-to-Face Environmental Science Course From 2009 to 2016. *Frontiers in Computer Science, 1*, 472525. https://doi.org/10.3389/fcomp.2019.00007

Shuran Zhang - Norlizah Che Hassan - Roshafiza Binti Hassan - Xiao Huang - Yujie Long - Cuiyun Wang / Eurasian Journal of Educational Research 111 (2024) 308-324

323

Rahmat, A. (2022). A Comparative Analysis Between Online and Offline English Learning in Covid-19 Pandemic Era: Students' View. *J-SHMIC: Journal of English for Academic, 9*(2), 39-54. https://doi.org/10.25299/jshmic.2022.vol9(2).9998

Reiss, R. D., Thomas, M., & Reiss, R. D. (1997). *Statistical Analysis of Extreme Values* (Vol. 2). Basel: Birkhäuser.

Salsabila Gita, N., & Utami, W. (2021). Students' Perceptions About the Online Listening Courses During the Covid-19 Pandemic. *Celtic : A Journal of Culture, English Language Teaching, Literature and Linguistics, 8*(1), 126-139. https://doi.org/10.22219/celtic.v8i1.16607

Samovar, L. A., & Porter, R. E. (1991). *Communication between Cultures*. Wadworth Publishing Company.

Schleeter, G. D., Slate, J. R., Moore, G. W., & Lunenburg, F. C. (2019). Gender Differences in the Reading Performance of Texas Grade 3 English Language Learners: A Multiyear, Statewide Investigation. *Asian Journal of Interdisciplinary Research, 2*(4), 141-155. https://doi.org/10.34256/ajir19413

Shachar, M., & Neumann, Y. (2003). Differences Between Traditional and Distance Education Academic Performances: A Meta-Analytic Approach. *International Review of Research in Open and Distributed Learning, 4*(2), 1-20. https://doi.org/10.19173/irrodl.v4i2.153

Sheerah, H. A. H., Yadav, M. S., Allah, M. A. E. F., & Abdin, G. J. (2022). Exploring Teachers and Students' Perceptions towards Emergency Online Learning Intensive English Writing Course during COVID-19 Pandemic. *Arab World English Journal (AWEJ) 2nd Special Issue on Covid 19 Challenges,* (2), 64-84. https://doi.org/10.24093/awej/covid2.5

Shen, Q., Chung, J. K. H., Challis, D., & Cheung, R. C. T. (2007). A comparative study of student performance in traditional mode and online mode of learning. *Computer Applications in Engineering Education, 15*(1), 30-40. https://doi.org/10.1002/cae.20092

Sipahutar, D. M. (2023). Comparison of Mathematics Learning Outcomes Online and Offline Students at Senior STIKES Medan Using Independent T-Test. *Edunesia : Jurnal Ilmiah Pendidikan, 4*(2), 607-618. https://doi.org/10.51276/edu.v4i2.413

Spitzer, M. W. H., & Musslick, S. (2021). Academic performance of K-12 students in an online-learning environment for mathematics increased during the shutdown of schools in wake of the COVID-19 pandemic. *PloS One, 16*(8), e0255629. https://doi.org/10.1371/journal.pone.0255629

Vargas-Quesada, B., Arroyo-Machado, W., Muñoz-Écija, T., & Chinchilla-Rodríguez, Z. (2023). Science overlay maps: A tribute to Loet Leydesdorff. *Profesional de la información, 32*(7), e320705. https://doi.org/10.3145/epi.2023.dic.05

Wang, G. Y. (2008). *A Study on the Current Situation of Academic Procrastination of Middle School Students, Its Causes and Correlation with Academic Achievement*. Nanjing Normal University. Nanjing, Jiangsu Province.

Wang, W. Y. (1998). Concepts, Strategies, and English Vocabulary Memory. *Foreign Language Teaching and Research,* (01), 49-54+80.

Wang, Y., Ma, J., Kremer, G. E., & Jackson, K. L. (2019). An investigation of effectiveness differences between in-class and online learning: an engineering drawing case study. *International Journal on Interactive Design and Manufacturing (IJIDeM), 13*(1), 89-98. https://doi.org/10.1007/s12008-018-0510-8

Wang, Z. (2021). Children's Gender Stereotype Threat and Its Interventions. *Advances in Psychological Science, 29*(2), 276-285. https://doi.org/10.3724/SP.J.1042.2021.00276

Shuran Zhang - Norlizah Che Hassan - Roshafiza Binti Hassan - Xiao Huang - Yujie Long - Cuiyun Wang / Eurasian Journal of Educational Research 111 (2024) 308-324

324

Wieser, D., & Seeler, J.-M. (2018). Online, Not Distance Education: The Merits of Collaborative Learning in Online Education. In A. Altmann, B. Ebersberger, C. Mössenlechner, & D. Wieser (Eds.), *The Disruptive Power of Online Education* (pp. 125-146). Emerald Publishing Limited. https://doi.org/10.1108/978-1-78754-325-620181008

Wu, G. (2020). Educational Technology as a Signpost for Learning in the Age of Risk: Dilemmas and Breakthroughs. *Open Education Research, 24*(03), 11-25. https://doi.org/10.13966/j.cnki.kfjyyj.2020.03.002

Xu, D., & Jaggars, S. S. (2011). The Effectiveness of Distance Education across Virginia's Community Colleges: Evidence from Introductory College-Level Math and English Courses. *Educational Evaluation and Policy Analysis, 33*(3), 360-377. https://doi.org/10.3102/0162373711413814

Zakarneh, B. M. (2018). Effectiveness of e-learning mode for teaching English language in Arab universities. *International Journal of Applied Linguistics and English Literature, 7*(7), 171-181. https://doi.org/10.7575/aiac.ijalel.v.7n.7p.171

Zhao, Y., Lei, J., Lai, B. Y. C., & Tan, H. S. (2005). What Makes the Difference? A Practical Analysis of Research on the Effectiveness of Distance Education. *Teachers College Record, 107*(8), 1836-1884. https://doi.org/10.1111/j.1467-9620.2005.00544.x