Article

Chinese University Students' Experience of WeChat-Based English-Language Vocabulary Learning

Fan Li [1,*], **Si Fan** [1], **Yanjun Wang** [2] **and Jinjin Lu** [1]

1   School of Education, College of Arts, Law and Education, University of Tasmania, Newnham Campus, Launceston, TAS 7248, Australia; si.fan@utas.edu.au (S.F.); jinjin.Lu@utas.edu.au (J.L.)
2   School of Humanities, College of Arts, Law and Education, University of Tasmania, Newnham Campus, Launceston, TAS 7248, Australia; yanjun.wang@utas.edu.au
*   Correspondence: f.li@utas.edu.au

Abstract: The outbreak of COVID-19 worldwide in 2020 has posed tremendous challenges to higher education globally. Teaching English as a foreign language (TEFL) is among the many areas affected by the pandemic. The unexpected transition to online teaching has increased challenges for improving and/or retaining students' language proficiency. WeChat, a popular social application in China, was widely used for TEFL at Chinese universities before COVID-19. However, it remains unclear whether the use of WeChat can facilitate Chinese university students' English-language lexical proficiency during the pandemic. To fill this gap, the aim of the present study was two-fold: (1) it initially explored the relationship between the variables including students' academic years, genders, and academic faculties/disciplines, and their lexical proficiency; and (2) it evaluated the effectiveness of a WeChat-assisted lexical learning (WALL) program in facilitating learning outcomes of English-language vocabulary. One hundred and thirty-three students at a university in Northern China participated in the WALL program for three weeks. As the results indicated, the independent variables had no correlation with the students' lexical proficiency. More importantly, the students had a decline in the test scores after using the program, compared to their initial test scores. Moreover, the difference was reported to be medium. The findings further proposed questions on applying WeChat to vocabulary teaching in a large-scaled transition. The study is expected to provide insights for tertiary institutions, language practitioners, and student stakeholders to troubleshoot the potential problems regarding implementing WeChat-based TEFL pedagogies.

Keywords: mobile-assisted language learning; English-language vocabulary learning; WeChat; Chinese higher education

1. Introduction

The COVID-19 pandemic has caused massive disruptions to educational institutions worldwide. With the advent of COVID-19 control measures, China mandated the nation-wide school closures at the end of February 2020 [1]. An emergency policy launched by the Chinese government—"Suspending Classes without Stopping Learning"—shifted higher education online [2]. This initiative minimized the impact of the pandemic on education [3] and ensured that learning was not disrupted at any point during the lockdown [4]. The online or web-based teaching, served as a substitute for the traditional way of classroom teaching, had been the mainstream mode of delivering teaching since late February 2020 [5]. Due to China's successful control measures, by June 2020, a large number of university students were able to return to campuses, resuming their studies in a blended mode [6].

Mobile learning (m-learning) refers to the phenomenon that mobile devices are applied for learning purposes [7]. As an integration of m-learning and language learning, mobile-assisted language learning (MALL) assists or enhances language learning in formal and/or informal environments, using handheld mobile devices, such as mobile phones or smartphones [8]. In spite of the widely recognized advantages of using MALL, there were

reports on unequal learning performances of some university students during the transition [9]. This calls for an investigation on MALL in the current global crisis context [10]. To address this gap, the present study aims to explore the factors influencing the lexical learning outcomes in the transition. It also evaluates the effectiveness of a WeChat-based MALL program used for English-language vocabulary learning at a Chinese university during the pandemic.

Numerous challenges during the pandemic, including technology accessibility, Internet connectivity, socioeconomic status, institutional supports, and learner experience, have left MALL effects uncertain [11]. It also remains questionable about the adoption of MALL-based pedagogies in higher educational systems in developing countries, such as China [12]. This particular study thus initially unveils the potential factors that might have impacts on students' English-language vocabulary proficiency. It also unpacks the effectiveness of equipping Chinese university students with a WeChat-based MALL program for English-language vocabulary learning under current circumstances.

To address the research objectives mentioned above, the paper initially unpacks related work primarily on MALL effects and WeChat-based language learning/teaching practices at Chinese tertiary levels, especially during the pandemic. It then presents that a WeChat-based MALL program which was used as an intervention. Next, the paper clarifies that the quantitative data were collected by using two sets of English-language vocabulary tests. Additionally, the test results were analyzed by using the Mann–Whitney U tests and Kruskal–Wallis tests to determine the relationships between the students' tests scores and several independent variables, such as their academic years, genders, and academic faculties/disciplines. The paired-sample *t*-test was then used to find out the statistically different significance between the test scores. The effect size was applied to further measure the difference, if there truly was a significance reported. Finally, the paper sheds light on a number of potential reasons that attributed to the findings.

2. Related Work

Studies of MALL did not receive as much attention as its wide prevalence in academia during the period of COVID-19 [13]. Nevertheless, m-learning has been a preferred learning manner for online language teaching practices during the pandemic [14]. MALL studies at the higher educational level focused primarily on participants' perceptions. For example, [15] surveyed 100 students and 45 teachers' reactions to the transition to mobile Russian language learning. Likewise, Chinese students held supportive attitudes that MALL approaches have made learning easier in their university English course particularly during the lockdown [10]. However, these studies mainly explored students and educators' beliefs and opinions on transforming to MALL paradigms rather than investigating its practical effects. Other studies, such as [16] and [17], analyzed affordances of implementing QQ and WeChat as MALL tools for university English courses facing the transition. The authors, however, largely put forward pedagogical implications of designing and applying MALL-based TEFL models, instead of measuring the actual effectiveness. Another study evidenced that voluntary out-of-class MALL has enhanced outcomes of learning French language by self-regulated training and scaffolding [18]. However, the study did not clarify whether the experiments were conducted in the transition and did not calculate the effect size of the treatment.

Shown as above, evaluating actual MALL effects on TEFL considering the pandemic situation was somehow overlooked, especially in the Chinese university context. For example, one most recent study reported mobile multimedia tools, including laptops or tablets, had positive impacts on Chinese university students' English learning outcomes under current situations [19]. However, the findings did not specify the facilitated language teaching area(s). As the most researched language-teaching area in MALL studies [20], mediated vocabulary learning outcomes have been investigated by using quantitative, qualitative, and/or mixed-methods approaches in a wealth of studies in the literature before the COVID-19 period [21]. Nevertheless, evaluating mobile-assisted vocabulary

learning effects in Chinese universities during the transition received seemingly little emphasis. Moreover, little empirical evidence scrutinized the effectiveness of specific media/platforms on MALL, such as WeChat.

WeChat has gained a high popularity in Chinese universities during the pandemic [22]. However, little research has explored using WeChat for university MALL practices amid the pandemic. One existing evidence is WeChat has developed students' enjoyment through emotion regulation in online collaborative English writing learning activities [23]. However, MALL on lexical proficiency has been sparse in the transition. Ref. [24] claimed, factors including user groups, academic contexts, and applications in use, largely attributed to successful MALL practices, particularly in vocabulary learning/teaching. Therefore, when evaluating the MALL effects, it is necessary to consider the factors mentioned above, because they potentially led to students' uneven MALL outcomes during the pandemic [25].

3. Research Aim and Questions

As previously discussed, the effectiveness of applying WeChat to vocabulary learning in the chosen context remains under-researched. To bridge this gap, 133 students at a Northern Chinese university participated in a WeChat-assisted lexical learning (WALL) program for a period of three weeks. The study addresses the following research questions (RQs):

RQ One: Do students' academic years, genders, and academic faculties/disciplines have impacts on their English-language vocabulary learning outcomes?

RQ Two: How well does the WALL program assist students to achieve their English-language vocabulary learning outcomes?

4. Research Method

4.1. Research Design

The reported study is under a larger scope PhD project. The research method underpinning the PhD project is a mixed-methods research design, including open-ended questionnaires, tests, and semi-structured interviews. This reported study was designed as the pilot study. It primarily focused on the quantitative data collected through two sets of vocabulary test papers, These tests included a Diagnostic Test administered before the WALL program and a Follow-Up Test administered afterwards. Differences between students' test scores collected before and after using the program were examined. The effect size was also calculated to determine whether the difference was large, medium, or small. The program duration was about three weeks, from 24 May to 21 June. Such a length of duration was decided, due to the fact that intervention or treatment shorter than one month was conducive to learners' improvement [26]. For longer programs, fatigue might appear [27], due to the novelty effect [28].

4.2. Participants

A purposive sampling strategy was applied to select the subject academic institution, sample faculties/disciplines, and sample participants in this study. Such a method was efficient in terms of time and effort saving, and suitable for a case study, as well as well-serving the research objectives [29]. Firstly, the university where the study was conducted is a key and comprehensive academic university in Northern China. University students were purposively chosen as the sample, because they are physically and mentally mature enough and are able to manage and discipline their learning tasks and pace [30]. Learning in universities is more independent than at other educational levels [31]. Besides, since this study was conducted via the WeChat public account on mobile phones, China has a large population of mobile phone users at universities [32]. Secondly, the selection of faculties/disciplines was also purposive. Among the total 24 faculties/disciplines, four faculties/disciplines were purposively selected as the representation, including the School of Architecture, School of Chemistry, School of Information Technology, and School of Media and Communication. Thirdly, the Year 1 and Year 2 students from the four faculties/disciplines were purposively chosen, because non-English majors at most Chinese

universities usually finish studying English courses at the end of Year 2. Initially, 150 students were recruited in a nonrandomized way [33]. Regarding the recruitment process, an invitation letter was posted on the websites of the faculties/disciplines after receiving the approval. Students volunteered for this study were contacted by the researchers via emails. Potential participants in naturally-formed classes were chosen as a convenience sample without randomization [34]. They were then informed of more details regarding this study before they made the final decision on their voluntary participation.

4.3. Demographic Backgrounds of the Participants

Totally, 150 students took the two sets of vocabulary tests online. One hundred and thirty-three submitted and completed both tests. Details of the students' demographic background information are presented in Table 1.

Table 1. Students' background information.

Background Information	% (n/N)
Academic faculties/disciplines/schools	
• Architecture	29.3 (39/133)
• Chemistry	16.5 (22/133)
• Information Technology	36.1 (48/133)
• Media and Communication	18.0 (24/133)
Genders	
• Male	75.9 (101/133)
• Female	24.1 (32/133)
Academic years	
• Year 1	57.9 (77/133)
• Year 2	42.1 (56/133)

4.4. Research Instruments

The WeChat-based MALL WALL program

The WALL program was designed and developed by the authors for university English vocabulary learning, including learning content and materials, daily practice and drills, and additional learning resources. The delivered content consisted of texts, audios, and video clips, covering a wide range of knowledge regarding a particular location, such as natural landscape, wildlife, social life, lifestyles, architectures, and cultures. The program was delivered through the public account service of WeChat. Students received daily notifications from the public account via WeChat the APP on mobile phones.

The English-language vocabulary test papers

Two sets of English-language vocabulary tests were used as the measurement tool. They were administered and collected online. In Phase 1, the Diagnostic Test, circulated before the program, was used to identify the students' initial lexical knowledge. The test paper had 25 multiple-choice questions. Each question contained one target lexical item and three distractor lexical items as interferences. Eighty lexical items were words and 20 were phrases. The lexical items were all extracted from the latest formal test papers of the College English Test: Band-4 (CET-4) (Retrieved from http://cet.neea.edu.cn/html1/folder/1608/1178-1.htm accessed on 19 May 2020) and selected based on the English teaching syllabus at the university. For example, "*The college students in China are ____from smoking on campus because this will do them no good. A. discouraged B. observed C. obeyed D. obtained*". Another example was "*After talking for nearly ten hours, he ____to the government's pressure at last. A. expressed B. yielded C. decreased D. approved*". In Phase 2, the Follow-Up Test, administered after the program, was used to examine the students' mediated lexical

knowledge. The test paper was designed at the same way as the former one. The lexical items in this paper were all based on the learning materials and content delivered in the program. One example was as follows: "*The wombat is a large ____found only in Australia. A. carnivore B. arthropod C. marsupial D. reptile*". Another example was as follows: "*The town of Binalong Bay is ____the southern end of the beautiful Bay of Fires. A. next to B. situated at C. near D. far from*". All lexical items were carefully selected according to the lexicon requirements in the *College English Curriculum Requirements* [35], compulsory active lexical items [36], and word frequencies [37]. Moreover, three experienced English-language teachers from the university were consulted to check all the target lexical items online before the program commenced. The practical relevance and difficulty levels of the lexical items were also verified.

Vocabulary tests were used as the test instrument to measure students' language proficiency and served as a vocabulary learning outcome indicator in this study for the following reasons. Firstly, vocabulary is the foundation of any language, and vocabulary education is a vital link in the chain of language acquisition [38]. Lexical ability is regarded more useful in English learning and teaching objectives [39] and plays a critical role in students' future careers in China [40]. Secondly, the glossary in the *College English Curriculum Requirements* is taken as a testing standard for lexical knowledge and as the criterion and norm reference for university English-language lexicons [41]. Thirdly, tests, as a comparatively easy and labor-saving way, can measure learners' vocabulary skills efficiently and accurately [42].

4.5. Data Analysis

The collected data were analyzed by using the Statistical Packages for Social Science (SPSS) version 26.0. Mean values and standard deviation of the students' test scores were firstly calculated for descriptive statistical analyses. Next, Mann–Whitney U tests were conducted to determine whether there were statistically significant differences between the students' test scores and independent variables, including academic years and genders. Then, Kruskal–Wallis tests were applied to examine whether significant differences existed between their test scores and academic faculties/disciplines. Finally, a paired-sample *t*-test was adopted to explore statistically significant differences between the two sets of test scores, namely the Diagnostic Test and the Follow-Up Test. If the significance value of the paired-sample *t*-test is smaller than 0.05, it shows that a significant difference is found between the students' scores on the two different tests. This indicated the WALL program has had an impact on the students' lexical proficiency. Otherwise, it can be said that no significance existed between the two sets of the test scores. Hence, the WALL program had no impact on the students' lexical proficiency. Subsequently, the effect size was calculated to measure whether the difference was large (d = 0.80), medium (d = 0.50), or small (d = 0.20), using Cohen's d [43], if there is any significance.

5. Results
5.1. Findings to Research Question One
5.1.1. Descriptive Analysis of the Test Results

The descriptive results of the two tests are first presented in this section. Table 2 shows the results of the Diagnostic Test and the Follow-Up Test, respectively. It can be seen that the mean value of the Diagnostic Test was 44.32, while the mean value of the Follow-Up Test was 31.10. The mean values above showed that overall, the students failed both tests, since 60 points (out of 100) is generally regarded as a passing score in China. A lower extent of statistical dispersions was also found on the individual participants' test scores on the Follow-Up English Test (standard deviation = 14.147) than the ones on the Diagnostic Test (standard deviation = 22.835). It can be claimed that despite a lower variation existed among the students' Follow-Up Test scores, their test performance was less satisfying as a whole. Additionally, according to the frequencies of the test scores from Table 3, 28.6% (N = 38) of the 133 students successfully passed the Diagnostic Test by scoring greater than

or equal to 60 points, while only 3.8% (N = 5) successfully passed the Follow-Up Test. A decline was found in the students' vocabulary proficiency test scores. That is to say, the students scored less after using the program than before they used it.

Table 2. Descriptive analysis of the two test scores.

	N	Min	Max	Mean	Standard Deviation
Diagnostic Test	133	14	94	44.32	22.835
Follow-up Test	133	8	96	31.10	14.147

Table 3. Frequencies of the two test scores.

	Diagnostic Test		Follow-Up Test	
Test Scores (Full Score: 100 Points)	**Frequencies (N = 133)**	**Percent (100%)**	**Frequencies (N = 133)**	**Percent (100%)**
91–100	4	3	2	1.5
81–90	9	6.8	0	0
71–80	13	9.8	1	0.8
60–70	12	9.0	2	1.5
Below 60 *	95	71.4	128	96.2

* Below 60 means failing the test.

5.1.2. Analysis of the Two sets of Test Scores by Independent Variables

Analysis of the Test Scores by Academic Years

Firstly, the Year 2 students had a higher average score on the Diagnostic Test (mean value = 47.11) than the Year 1 cohort (mean value = 42.30). It was because the Year 2 students, in most cases, had more exposure to vocabulary learning and teaching than their Year 1 counterparts. However, neither group scored greater than or equal to 60 points on average. Next, a Mann–Whitney U test indicated no statistically significant difference between the students' Diagnostic Test scores and their academic years, since the p-value (0.372) was larger than 0.05. Secondly, the Year 1 students had better performances on the Follow-Up Test (mean value = 32.26) than the Year 2 group (mean value = 29.50). The Year 1 students could be more likely to follow the learning requirements. They could possibly be more hardworking and maintain better learning effectiveness, since they had not graduated from high schools for long. However, neither group scored greater than or equal to 60 points. Then, a Mann–Whitney U test presented no statistically significant difference was found between the students' Follow-Up Test scores and their academic years, since p-value = 0.098, which was larger than 0.05. It could be stated that the students' academic years have had no impacts on the two sets of test scores.

Analysis of the Test Scores by Genders

Firstly, the female students had a much higher score on the Diagnostic Test (mean value = 64.56) than their male counterpart (mean value = 43.61). The female students passed the test on average, while the male students did not do well. Female students generally showed better language learning outcomes and greater devotion to schooling. Next, a Mann–Whitney U test found no statistically significant difference existed between the students' Diagnostic Test scores and their genders, since the p-value (0.457) was larger than 0.05. Secondly, the female students had a much higher score on the Follow-Up Test (mean value = 31.33) than the male students (mean value = 30.37). Female students, again, contributed slightly better test performances. Both groups, however, did not pass the test. Mann–Whitney U test then showed no statistically significant difference between the students' Follow-Up Test scores and their genders, since the p-value (0. 769) was larger

than 0.05. It can be observed that the students' genders have had no impacts on the two sets of test scores.

Analysis of the Test Scores by Academic Faculties/Disciplines

Firstly, the result showed the students from the School of Media and Communication had the highest mean score on the Diagnostic Test (mean value = 46.29) among the four academic faculties/disciplines, followed by the ones from the School of Chemistry (mean value = 46.18) and the School of Information Technology (mean value = 44.92). The students from the School of Architecture had the lowest mean score (mean value = 41.33). The reason was possibly because arts students generally had better performances in English study than science and engineering students. However, all four groups failed the test, since their mean scores were lower than 60 points. The Kruskal–Wallis test result on scores by academic faculties/disciplines indicated the students from the School of Information Technology had the highest test score (mean rank = 70.27). However, the students from the School of Architecture scored the lowest (mean rank = 61.71). The result indicated no correlation between the students' the Diagnostic Test scores and their academic faculties/disciplines ($x2 = 1.151$, df = 3, *p*-value = 0.765 > 0.05). Secondly, the students from the School of Information Technology scored the highest on the Follow-Up Test (mean value = 33.75) among the four academic faculties/disciplines, followed by the ones from the School of Chemistry (mean value = 31.82) and the School of Media and Communication (mean value = 9.916). However, the students from the School of Architecture had the lowest mean score (mean value = 28.51). The reason why Computer Science had better test scores could be that they were more familiar with and adept at using mobile technologies. They could thus take advantages of the program to its fullest. However, all students failed the test again since their mean scores were lower than 60 points. The result of the Kruskal–Wallis test on the scores by academic faculties/disciplines showed the students from the School of Information Technology had the highest test score (mean rank = 71.35). However, the students from the School of Architecture scored the lowest (mean rank = 61.62). There was no correlation between the students' Follow-Up Test scores and their academic faculties/disciplines ($x2 = 1.401$, df = 3, *p*-value = 0.705 > 0.05). It could be concluded that the students' academic faculties/disciplines have had no impacts on the two sets of test scores.

5.2. Findings to Research Question Two

Paired-Sample *t*-Test Result of the Two Sets of the Test Scores

A paired-sample *t*-test was conducted to compare and analyze the 133 students' scores on each vocabulary proficiency test to verify whether a statistically significant difference existed between the two test scores. According to the paired-sample *t*-test result (as shown in Table 4 below), a statistically significant difference was found between the students' two sets of test scores, since the significance value = 0.000, which was smaller than 0.05. That is, the 133 students did perform statistically significantly on the two sets of tests. Additionally, the effect size was calculated to further measure the WALL program effect on the test scores. Cohen's d indicated the effect size magnitudes, namely d = 0.20 for a small effect, d = 0.50 for a medium effect, and d = 0.80 for a large effect [43]. It was reported the WALL program had a medium effect size of 0.50. That is, the WALL program has brought about medium differences to the students' test scores.

Table 4. Paired-sample *t*-test result of the test scores.

	Paired Differences		t	df	Sig. (2-Tailed)
	Mean	**Standard Deviation**			
Diagnostic Test Follow-Up Test	13.226	26.460	5.764	132	0.000

6. Discussions

The previous results initially indicated a number of potential influential factors/variables, including the students' academic years, genders, and academic faculties/disciplines, had no correlations with the test scores. More importantly, the students had statistically significantly different test scores. However, it was discouraging to see the decline in the test scores after the delivery of the WALL program. Therefore, it can be argued that the program was the only variable that has had negative impacts on the 133 students' test performances. That is to say, the students did not improve or retain their lexical proficiency, despite they had used the program for the period of three weeks during the pandemic. Contrary to the findings of Reference [24] which reported MALL positive effects by synthesizing 80 latest worldwide publications before the pandemic, the results of this particular study showed the MALL approach did not improve or retain students' lexical proficiency in the transition. Negative results, however, were not unusual according to previous evidence [44]. Several reasons could possibly explain the unsuccessful intervention of the WALL program:

First, it was a challenge for a considerable number of students to embrace the new MALL approach. It was true that, before COVID-19, MALL pedagogies were welcomed [45], and mobile phones or smartphones were the most widespread mobile device for MALL practices [46]. However, according to the statistics shown on the WeChat public account, the students in this study hesitated to engage with the program. Explicitly, not all the students have viewed the learning materials and content, performed the daily practice and drills, or read the additional learning resources. This happened probably because they cast doubts on integrating WeChat with the regular teaching syllabus in the face of the unexpected transition. It was consistent with [6]'s findings that around half of the participants remained uncertain about effects of the new learning approach. Students would have a strong intention of conducting MALL activities, if only they perceived them to be pronouncedly useful [10].

Second, limited technical skills and knowledge prohibited students in using MALL tools. Owing to the large population of mobile phone users in China, it is hard to see a student without a mobile phone on university campuses [47]. The generation of digital nativities was familiar with and not irritated by using mobile devices for multifaceted purposes [48]. In addition, most Chinese university students would like to use mobile technologies for learning activities during the pandemic [49]. The subject university was equipped with on-campus high-speed network and stable Internet connectivity. Most of all, the program was designed while considering the costs of mobile data and challenges of network speed. However, certain students in this study were not familiar with fully applying the WeChat public account for vocabulary learning. They were possibly uncertain and confused about using the program to the fullest, ranging from reading learning materials in multiple forms, doing daily practice and drills, to viewing more learning resources. It could also be overwhelming and complex for some students to download texts, audios, and video clips from the program for future use. As mentioned previously, the vocabulary-learning content and materials were purposively designed in various forms. Language learning, however, could be interfered by redundancy and working memory load if students are exposed to the concurrent delivery of same learning information in different modes [50].

Third, students had difficulty concentrating on MALL activities. The students in this study could possibly devote to social chats when logging in the public account via WeChat. They would then get immersed in chatting with their friends rather than learning with the program. This happened since WeChat is a social application in which messages pop up automatically. Meanwhile, notifications and new information from other subscribed public accounts could also distract them from engaging with the program. When scrolling down the list of subscribed public accounts, the students' attention could be drawn by other appealing information, such as games. Since the learning activities in this study were carried out in an informal learning setting, the students were likely to be disturbed by

any trivial matters. Undoubtedly, environments free from distractions were conducive to enhancing MALL outcomes [49].

Fourth, students had difficulties participating in MALL activities. As an indispensable element in educational practices, teachers had crucial impacts on students' learning performances [51]. In this study, the students experienced reduced engagement in the lexical learning activities when using the program in an informal/non-classroom learning setting. It happened because of the absence of language teachers who often supervise or remind students of learning tasks performed in formal/classroom learning environments. Students mostly conducted autonomous learning or self-directed learning during the pandemic [52]. That is, it required self-regulation and self-discipline from the students when using the program in absence of teacher-led instruction. However, as the statistics the public account presented, numerous students would not like to use the program autonomously until their language teachers asked them to.

Fifth, students' negative emotions could possibly have hindered the MALL effects. Students, when learning online, were likely prone to negative emotions, such as frustration, anxiety, learning slackness, and weariness [14]. In this study, the students mostly had to learn the lexical items and perform the learning activities on the program on their own. Therefore, they would possibly feel disconnected and distant resulting from limited interaction due to lack of physical proximity and isolation with peers. Such issues are common in informal online learning practices [53]. Moreover, the students had weaker intention and motivation in learning vocabulary by using the program, as their excitement for participating in MALL activities could be temporary [54]. Therefore, the students in this study would possibly have less satisfactory results when their enthusiasm decreased during the period of program delivery.

7. Limitations and Suggestions for Future Research

The study has several limitations. Firstly, considering the small sample size of the case study, the findings could be less representative regarding different situations at other Chinese universities. The authors would suggest that further studies should have a larger sample size and a wider geographical range, because it would be beneficial to provide a comprehensive picture of the researched topic. Since the present project was conducted in an academic university context, investigations in vocational colleges were also suggested. Secondly, since the university in this study allowed the students to return to the campus before the intervention started, the students used the program for lexical learning in a mixed learning mode. They might have conducted learning activities, using the program either in or out of the classroom. Findings could thus be different if the setting was completely formal or informal. The authors call for future studies that should consider different physical locations/environments under these circumstances. Thirdly, the present study selected the participants by using a nonrandomized sampling method. The authors would suggest that future studies can apply randomization for the participant recruitment to verify the research findings. Fourthly, the present paper primarily investigated the quantitative findings, since it was under a larger PhD project. The students' opinions and perceptions of using the program for MALL practices during the pandemic were explored in a separate paper of the authors, using a mixed-methods research design, including open-ended questionnaires and semi-structured interviews. The paper will be drafted and submitted soon. Fifthly, considering the fact that MALL approaches have been widely used for a broad range of language skills [55], the present study that has focused specifically on vocabulary calls for future studies to investigate different language teaching areas. Lastly, the intervention duration was around one month, owing to the time conflict with the final-exam week at the university. Since intervention durations have been found to impact the MALL effectiveness [56], future studies are expected to conduct a longitudinal empirical study with longer intervention durations.

8. Conclusions

The present study evaluated the effectiveness of a WeChat-based program in facilitating students' vocabulary learning outcomes at a Northern Chinese university. Two sets of English-language vocabulary tests were used to measure the students' lexical proficiency. The findings initially indicated that the independent variables, such as the students' academic years, genders, and academic faculties/disciplines, had no impact on their vocabulary test performances. Moreover, it was found that the students had a decline in their test scores after using the WALL program for a period approximately three weeks. The program has had a medium effect on the difference of the test scores. It can be stated that the effectiveness of MALL in the Chinese higher-education context remains debatable amid the pandemic.

References

1. UNESCO. 290 Million Students out of School Due to COVID-19: UNESCO Releases First Global Numbers and Mobilizes Response. 2020. Available online: https://en.unesco.org/news/290-million-students-out-school-due-covid-19-unesco-releases-first-global-numbers-and-mobilizes (accessed on 24 June 2021).
2. Ministry of Education of the People's Republic of China. An Announcement Regarding "Suspending Classes without Stopping Learning". 2020. Available online: http://www.moe.gov.cn/jyb_xwfb/gzdt_gzdt/s5987/202002/t20200212_420385.html (accessed on 24 June 2021).
3. Hebebci, M.T.; Bertiz, Y.; Alan, S. Investigation of views of students and teachers on distance education practices during the Coronavirus (COVID-19) Pandemic. *Int. J. Technol. Educ. Sci.* **2020**, *4*, 267–282. [CrossRef]
4. Zhang, W.; Wang, Y.; Yang, L.; Wang, C. Suspending classes without stopping learning: China's education emergency management policy in the COVID-19 outbreak. *J. Risk Financ. Manag.* **2020**, *13*, 55. [CrossRef]
5. Gao, L.X.; Zhang, L.J. Teacher learning in difficult times: Examining foreign language teachers' cognitions about online teaching to tide over COVID-19. *Front. Psychol.* **2020**, *11*, 1–14. [CrossRef] [PubMed]
6. Su, H. Educational assessment of the post-pandemic age: Chinese experiences and trends based on large-scale online learning. *Educ. Meas. Issues Pract.* **2020**, *39*, 37–40. [CrossRef]
7. Kukulska-Hulme, A.; Shield, L. An Overview of Mobile Assisted Language Learning: Can Mobile Devices Support Collaborative Practice in Speaking and Listening. Available online: https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.84.1398&rep=rep1&type=pdf (accessed on 24 June 2021).
8. Kukulska-Hulme, A.; Shield, L. An overview of mobile assisted language learning: From content delivery to supported collaboration and interaction. *Recall* **2008**, *20*, 271–289. [CrossRef]
9. Yang, R. China's higher education during the COVID-19 pandemic: Some preliminary observations. *High. Educ. Res. Dev.* **2020**, *39*, 1317–1321. [CrossRef]
10. Guo, J.Y.; Huang, F.; Lou, Y.Q.; Chen, S.M. Students' perceptions of using mobile technologies in informal English learning during the COVID-19 epidemic: A study in Chinese rural secondary schools. *J. Pedagog. Res.* **2020**, *4*, 475–483. [CrossRef]
11. Moser, K.M.; Wei, T.; Brenner, D. Remote teaching during COVID-19: Implications from a national survey of language educators. *System* **2021**, *97*, 102431. [CrossRef]
12. Khan, A.I.; Al-Shihi, H.; Al-Khanjari, Z.A.; Sarrab, M. Mobile Learning (M-Learning) adoption in the Middle East: Lessons learned from the educationally advanced countries. *Telemat. Inform.* **2015**, *32*, 909–920. [CrossRef]
13. Naciri, A.; Baba, M.A.; Achbani, A.; Kharbach, A. Mobile learning in Higher education: Unavoidable alternative during COVID-19. *Aquademia* **2020**, *4*, ep20016. [CrossRef]

14. Zhou, Z. On the lesson design of online college English class during the COVID-19 pandemic. *Theory Pract. Lang. Stud.* **2020**, *10*, 1484–1488. [CrossRef]
15. Kolesova, D.V.; Moskovkin, L.V.; Popova, T.I. Urgent transition to group online foreign language instruction: Problems and solutions. *Electron. J. e-Learn.* **2021**, *19*, 21–41. [CrossRef]
16. Peng, S.X.; Luo, J. SWOT analysis of foreign language teaching online under the background of the "COVID-19 Epidemic": A case study of Kaili University. *J. Kaili Univ.* **2020**, *38*, 81–85.
17. Song, J.Y. Investigating a m-learning model for college English from the perspective of "AI+Education". *J. Kaifeng Vocat. Coll. Cult. Art* **2020**, *40*, 87–88.
18. Botero, G.A.G.; BOTERO, M.A.; Zhu, C.; Questier, F. Complementing in-class language learning with voluntary out-of-class MALL. Does training in self-regulation and scaffolding make a difference? *Comput. Assist. Lang. Learn.* **2019**, *32*, 1–16. [CrossRef]
19. Yu, Z.G. The effects of gender, educational level, and personality on online learning outcomes during the COVID-19 pandemic. *Int. J. Educ. Technol. High. Educ.* **2021**, *18*, 1–17. [CrossRef]
20. Burston, J. Twenty years of MALL project implementation: A meta-analysis of learning outcomes. *Recall* **2015**, *27*, 4–20. [CrossRef]
21. Lin, J.J.; Lin, H.F. Mobile-assisted ESL/EFL vocabulary learning: A systematic review and meta-analysis. *Comput. Assist. Lang. Learn.* **2019**, *32*, 878–919. [CrossRef]
22. Liu, Y.F. Investigating College Students' English Online Learning Autonomy during the COVID-19 Pandemic. Master's Thesis, Hebei University, Baoding, China.
23. Zhang, Z.; Liu, T.; Lee, C.B. Language learners' enjoyment and emotion regulation in online collaborative learning. *System* **2021**, *98*, 102478. [CrossRef]
24. Chen, Z.Z.; Chen, W.C.; Jia, J.Y.; An, H.L. The effects of using mobile devices on language learning: A meta-analysis. *Educ. Technol. Res. Dev.* **2020**, *68*, 1769–1789. [CrossRef]
25. Al-Emran, M. Mobile Learning during the Era of COVID-19. *Rev. Virtual Univ. Católica Del Norte* **2020**, *61*, 1–2. Available online: https://revistavirtual.ucn.edu.co/index.php/RevistaUCN/article/view/1195 (accessed on 24 June 2021). [CrossRef]
26. Chiu, Y.H. Computer-assisted second language vocabulary instruction: A meta-analysis. *Br. J. Educ. Technol.* **2013**, *44*, E52–E56. [CrossRef]
27. Peng, H.; Jager, S.; Lowie, W. Narrative review and meta-analysis of MALL research on L2 skills. *Recall* **2020**, *33*, 1–18.
28. Stockwell, G.; Hubbard, P. Some emerging principles for mobile-assisted language learning. *Int. Res. Found. Engl. Lang. Educ.* **2013**, *2013*, 1–15. Available online: http://www.tirfonline.org/english-in-the-workforce/mobile-assisted-language-learning (accessed on 24 June 2021).
29. Qin, X.Q. *Questionnaires in Foreign Language Teaching*; Foreign Language Teaching and Research Press: Beijing, China, 2009.
30. Klassen, J.; Vogel, D. New issues arising from e-education. In *Web-Based Education: Learning from Experience*; Aggarwal, A., Ed.; IRM Press: Hershey, PA, USA, 2003.
31. Fan, S. Significance of the Web as a Learning Resource in an Australian University Context. Ph.D. Thesis, University of Tasmania, Launceston, Tasmania, 2011.
32. CNNIC. The 42nd China Statistical Report on Internet Development. 2018. Available online: http://www.100ec.cn/home/detail--6467393.html (accessed on 24 June 2021).
33. Mahendran, R.; Tan, J.Y.S.; Griva, K.; Lim, H.A.; Ng, H.Y.; Chua, J.; Kua, E.H. A pilot, quasi-experimental, mixed methods investigation into the efficacy of a group psychotherapy intervention for caregivers of outpatients with cancer: The COPE study protocol. *BMJ Open* **2015**, *5*, 1–8. [CrossRef]
34. Creswell, J.W. *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*, 4th ed.; Sage: Los Angeles, CA, USA, 2014.
35. Ministry of Education of the People's Republic of China. *College English Curriculum Requirements*; Shanghai Foreign Language Education Press: Shanghai, China, 2007.
36. Wang, X.Y. Patterns and Causes of Attrition of English as a Foreign Language. Ph.D. Thesis, Shandong University, Jinan, China, 2011.
37. Ellis, N. The associative-cognitive CREED. In *Theories in Second Language Acquisition*; VanPatten, B., Williams, J., Eds.; Erlbaum: Mahwah, NJ, USA, 2007; pp. 77–96.
38. Jiang, T.X. An empirical study on English vocabulary attrition of college non-English majors. *Valuab. Proj.* **2014**, *16*, 258–260.
39. Niu, L.W. An investigation of demotivation in English learning of English majors. *J. Hubei Radio Telev. Univ.* **2016**, *36*, 32–36.
40. Li, F.; Liu, X.J. Qualitative research on lexical fossilized errors in written language of the non-English majors. *J. Chengdu Univ. Technol. Soc. Sci.* **2016**, *2*, 134–140.
41. National Testing Board for Band-4. *Syllabus for College English test: Band Four*; Shandong Foreign Language Education Press: Shanghai, China, 2006.
42. Zou, S. Testing specifications and reading assessment: Designing the new TEM-4 reading testing specifications. *CELEA J.* **2007**, *30*, 3–15.
43. Cohen, J. *Statistical Power Analysis for the Behavioral Sciences*, 2nd ed.; Lawrence Erlbaum: Hillsdale, MI, USA, 1988.
44. Derakhshan, A.; Kaivanpanah, S. The impact of text-messaging on EFL freshmen's vocabulary learning. *Eur. Assoc. Comput. Assist. Lang. Learn.* **2011**, *19*, 30–39.
45. Gorjian, B. Teaching Vocabulary through Web-Based Language Learning (WBLL). *Approach. Procedia Technol.* **2012**, *1*, 334–339. [CrossRef]

46. Jarvis, H.; Achilleos, M. From computer assisted language learning (CALL) to mobile assisted language use (MALU). *TESLEJ* **2013**, *16*, 1–18.

47. Cui, Y.H. An Empirical Study of Mobile Learning of College English Vocabulary. Master's Thesis, Yanan University, Shaanxi, China, 2014.

48. Prensky, M. Digital natives, digital immigrants part 1. *Horizon* **2001**, *9*, 1–6. [CrossRef]

49. Xie, X.; Siau, K.; Nah, F.F.H. COVID-19 pandemic–online education in the new normal and the next normal. *J. Inf. Technol. Case Appl. Res.* **2020**, *22*, 175–187. [CrossRef]

50. Kalyuga, S.; Sweller, J. The redundancy principle in multimedia learning. In *The Cambridge Handbook of Multimedia Learning*, 2nd ed.; Cambridge University Press: New York, NY, USA, 2014; pp. 247–262.

51. Hew, K.F. Promoting engagement in online courses: What strategies can we learn from three highly rated MOOCS. *Br. J. Educ. Technol.* **2016**, *47*, 320–341. [CrossRef]

52. Zhao, N.; Zhou, X.; Liu, B.; Liu, W. Guiding teaching strategies with the education platform during the COVID-19 epidemic: Taking Guiyang No. 1 Middle School teaching practice as an example. *Sci. Insigt. Edu. Front.* **2020**, *5*, 531–539. [CrossRef]

53. Jung, Y.; Lee, J. Learning engagement and persistence in massive open online courses (MOOCS). *Comput. Educ.* **2018**, *122*, 9–22. [CrossRef]

54. Gay, G.; Stefanone, M.; Grace-Martin, M.; Hembrooke, H. The effects of wireless computing in collaborative learning environments. *Int. J. Hum.-Comput. Interact.* **2001**, *13*, 257–276. [CrossRef]

55. Viberg, O.; Grönlund, Å. Mobile assisted language learning: A literature review. In Proceedings of the 11th World Conference on Mobile and Contextual Learning, Helsinki, Finland, 16–18 October 2012.

56. Crompton, H.; Burke, D. The use of mobile learning in higher education: A systematic review. *Comput. Educ.* **2018**, *123*, 53–64. [CrossRef]