The Impact of Mobile Learning on the Effectiveness of English Teaching and Learning–A Meta-Analysis

MENG-LIN CHEN [ID]
Department of Translation and Interpretation Studies, Chang Jung Christian University, Guiren, Tainan 71101, Taiwan

e-mail: menglin@mail.cjcu.edu.tw

ABSTRACT Mobile devices have been widely used in both teaching and learning English as a second or foreign language. The impact of using various mobile devices on the learning outcome has thus been subject to an increasing number of educational experiments. However, the results of these experiments have not always been consistent or conclusive. In order to evaluate the effectiveness of mobile learning, it is necessary to employ a meta-analysis approach to synthesize results taken from previous experiments. This study examined 29 experiments, including both true and quasi, on the learning outcome of those learners who used mobile devices in their English learning process. Following *PRISMA* guidelines for conducting meta-analysis, this study has found that mobile learning has a highly positive effect on English learning outcomes with a weighted effect size of 0.893. The effect of mobile learning is subject to potential moderating variables, such as learners' levels, intervention time, language learning areas being targeted, hardware and software, and implementation setting. Taken as a whole, the results indicate that mobile learning is affording a change in English learning that does not undermine pre-existing models of English language pedagogy. Therefore, progressive exploration and implementation of mobile learning effectiveness appears to be of some importance to English teaching and learning stakeholders.

INDEX TERMS Effect, English learning, meta-analysis, mobile learning, quasi-experiment, true experiment.

I. INTRODUCTION

Mobile learning (M-learning) is a form of learning in which learners use mobile devices to access learning materials and resources and to participate in learning activities through wireless technology [1]. It is often referred to as Mobile Assisted Language Learning (MALL) in the context of second/foreign language teaching and learning [2]–[4]. With the rapid development of the Internet and mobile technology, new technologies are increasingly accessible to learners in most parts of the world, and M-learning has gradually emerged as a trend in learning at all levels [5]. It not only enhances students' learning interest and motivation [4], but also helps students obtain necessary technological literacy, which plays an important role for them to succeed in future workplace and lifestyle [6]. From a pedagogical perspective, M-learning has been considered as a more effective teaching tool than other forms of computer-assisted instruction [3] since it offers new

The associate editor coordinating the review of this manuscript and approving it for publication was Ali Shariq Imran [ID].

opportunities for students to actively engage in teaching and learning opportunities [6].

In the area of English and foreign language teaching and learning, M-learning has been used formally and informally for some time, and numerous studies have investigated the impact of mobile English learning using different approaches, qualitative, quantitative, meta-analytical etc. [2]–[4], [7]–[11]. Many true-experiment or quasi-experiment studies have provided important insight into the impact of information technology on Learning Performance (LP) [1], [12], [13]. The majority of such studies have claimed the significant and positive effect of mobile learning on English learning effectiveness, such as with vocabulary, oral language, listening, and retention. For example, in a pre-test/post-test experiment of non-equivalent groups (26 experimental and 31 control) from two middle school classes, the experimental group that used mobile applications were found to perform better than the control group that learned in a traditional teacher-led manner after an intervention period of 6 weeks [9]. In a quasi-experiment on

Bring Your Own Device (BYOD) for eighth grade middle school students to learn English [4], the experimental group (24 students) that used mobile devices for test-based learning had significantly higher learning outcomes in the post-test than the control group (22 students) settings, which used paper-based test sheets for learning after 6 sessions (45 minutes each) of intervention. In a random group experiment [14] with 116 participants from a university in central Taiwan, the experimental group was instructed to learn English phrases by taking pictures using their cell phones, while the control group was assigned online phrase reading activities for phrase learning; the former were found in the post-test to significantly outperform the latter with a significantly higher perception of English phrases. In a similar study of mobile phone-assisted learning, it was found that mobile phones could facilitate improvement in pronunciation with all participants providing positive feedback on the effectiveness of mobile application-assisted English learning [15].

However, some experimental and quasi-experimental studies claimed the negative or weak effect of M-learning on English language-learning effectiveness. For example, [16] developed a mobile immersion environment using a real-time WhatApp and conducted a 3-month pre-test/post-test non-equivalent group experiment with 45 grade 7 students in an elementary school. It was found that there was no significant difference between the experimental group and the control group in the post-test of vocabulary acquisition and use of high frequency verbs. Reference [17] found that students' familiarity with the real environment of M-learning was more important than the technology itself. In reality, not all students were familiar with the real environment leading to M-learning technology being unbeneficial for all students. Nevertheless, the total number of studies claiming the negative effects of M-learning was quite small. In fact, according to [1], among 164 studies in M-learning studies between 2003 and 2010, only 1% had reported a negative effect on the learning outcome.

In addition, the results of some non-experimental studies also showed inconsistent results. Some studies claimed positive effects of M-learning on English learning [11], [18]–[20]. For example, using questionnaires and interviews to analyze students with average computer and English language proficiency, it was found that M-learning was beneficial to students' academic writing and presentation skills [11]. Others claimed the negative or neutral effect of m-learning on English teaching and learning [21]–[23]. For example, a survey of 126 EFL students and 73 teachers showed that teachers and students only had a moderately positive attitude towards using mobile dictionaries in the learning process [22]. In addition, it was found that some teachers and students had opposing views on the effects of using mobile devices in teaching, with some teachers even prohibiting the use of mobile phones in the classroom for their distraction of student attention [23]. In her opinion [23], the use of digital tools in language classrooms lacked cultural background, and there were various problems related

with personal learning methods, teaching methods, the use of digital devices as entertainment tools, students' learning preferences, and motivations leading to the ineffectiveness of mobile devices.

Some researchers also conducted systematic reviews of previous literature related to the effectiveness of M-learning on language teaching and learning [1], [3], [6], [12], [13], [24]–[27]. They mainly discussed whether, and under what, circumstances involving M-learning was effective. Some of the reviews were based on a considerably large amount of literature. For example, [1] and [25] both looked at as many as 164 studies, and Persson and Nouri [3] reviewed 54 papers on Mobile Assisted Second Language Learning (MASLL). The majority of these reviews suggested that although M-learning was effective for English learning, the effectiveness was contextually dependent [5], [12], [13]. The success of M-learning relied on a host of contextual and situational factors including communication and feedback between teachers and students who used mobile devices in teaching and learning [13], students' technological skills, the development of learner communities, learners' attitudes towards M-learning, and learning content involved [5].

In addition, these studies also provided some useful insights. M-learning studies have been found to focus more on applications (apps) than on hardware such as iPads, iPhones, mobile phones, etc. [3], [24]. And of all the M-learning applications developed, games were the most effective [24]. While M-learning improved language learners' overall abilities including listening, speaking, reading, and writing in English [24], it was most effective for vocabulary development [3].

These reviews also pointed to some of the more significant challenges to existing research related to mobile language learning. First, the majority of studies in this area have been conducted on university student populations, and the research on other groups of learners has been insufficiently proportional [3], [13], [24]. Many of these studies proposed some newly-developed devices or applications, very few of which were commercially available, rendering the studies isolated in terms of technology and non-conducive to market promotion [13]. Another problem was that some apps were not always affordable for learners [5]. While pedagogical aspects of M-learning remained another common concern [3], [24], there was also a research methodological issue for many researchers in the studies being M-learning app developers and the instructors in the experiments. This might influence the validity of their research. For example, it was found that such studies in which the researchers coincidentally played the role of app developers and teachers usually reported larger M-learning effects than in other studies [6]. Furthermore, a number of important factors such as the motivation and attitude of both students and teachers have been identified as key to the success of M-learning that require greater attention [6], [12], [24].

Although some studies [1], [3], [24], [25] have included more than 100 research papers in their systematic review,

the meta-analysis approach was not followed. Without the information on the heterogeneity of subjects, interventions, and experimental designs, their results did not provide evidence of high-level strength to form definitive conclusions about the effectiveness of mobile devices on English language teaching and learning. For example, one study focused on statistical descriptions of the frequency of research methods (interviews, experiments, observations, and cases), research objectives, subjects, and device types for M-learning. Statistical descriptions of the findings were limited to positive results, negative results, and percentages of insignificant effects [25]. Another study used only manual coding [3]; Yet another study used strict inclusion criteria and only general statistics without consideration of the heterogeneity of challenges and concerns of M-learning as well as its strengths and weaknesses [24].

To the best of our knowledge, the only meta-analysis studies of M-learning were focused on either broader scopes beyond language learning or on more narrowed context of applying a particular M-learning technology [6], [12], [13], [26]. For example, one study of M-learning took place with various participants in K-12 (including English and mathematics instruction) [6]. Another one selected only 9 studies of EFL class instructions using mobile devices to look at their sample size, experimental design, and learning objectives (e.g. vocabulary, reading, writing, grammar, and retention) of using mobile devices [26].

In summary, although many studies have looked at English M-learning e.g. [3], [9], [26], [28], they did not always focus on the effectiveness of mobile English learning and their findings were inconsistent [11], [18], [19], [21]–[23]. This highlights the need to synthesize studies on the impact of English M-learning. According to [29], meta-analysis provides a sound approach to synthesize the existing evidence of mobile English learning. This study thus uses a meta-analysis approach to determine the effects of mobile English learning in the international context. In this way, the overall and specific effects of English M-learning will be investigated, providing researchers and educators with multiple perspectives.

II. METHODS

Meta-analysis is a statistical analysis method that integrates the results of several independent studies, with early applications used in the medical field. Effect size is an important value in the meta-analysis process, reflecting the strength of the relationship between two variables: intervention and effect. In this study, a meta-analytic approach was used to code experimental and quasi-experimental studies of mobile English learning. Based on the meta-analysis steps and methods proposed by [30], the mean effect size of the studies was first calculated and then followed by a calculation of the effects of other variables on the effect size to fully investigate the effectiveness of mobile English learning. Although Cohen's effect size d was commonly used as an indicator of effects in meta-analyses, it might produce some bias in the estimation of effects for small samples due to the variation in

sample sizes between the studies [31]. Therefore, Hedges' g unbiased effect size was used instead in this study.

Meta-analyses use fixed-effect or random-effect models for estimating the effect size based on the post hoc heterogeneity test. If the variance of Q value is not significant, the fixed-effect model is selected; Otherwise, the random-effect model is selected, and F value is used to determine the proportion of the true variance of the eigenvalues to the total variance [29]. The higher the degree of heterogeneity of the effect size, the more rational the random-effect model is chosen. In this study, Comprehensive Meta Antalysis (CMA) version 3.0 [32] was used to conduct the meta-analysis after values of means and standard deviations of pre-tests and post-tests information of the experiments were entered.

In addition, to ensure transparency, accuracy, and completeness, the meta-analysis conducted in our study is in accordance with the steps outlined in the *PRISMA* (Preferred Reporting Items for Systematic Reviews and Meta-Analysis) guidelines. The 11 important methodological issues numbered 5-15 in the method checklist of *PRISMA* Statement (2020) [33] were all addressed in the following sections.

A. SEARCH, SELECTION AND CODING OF EXPERIMENT SAMPLES

True experimental studies, in which subjects are completely randomized and variables are strictly controlled, include pre-test-post-test equivalent-group, post-test-only equivalent-groups, and randomized controlled trials [32]. However, the majority of experimental studies in the field of English teaching and learning are not true experiments, because in reality it is more difficult to randomly select research participants from the total population or to randomly group them. Researchers in this field tend to carry out quasi-experiments under real teaching situations with controlled experimental factors. In quasi-experiments, researchers use subject groups naturally formed, e.g. groups based on original classes, and intervene in routine English language teaching contexts. The most common types of quasi-experiments are pre-test-post-test non-equivalent-groups and counter balanced designs [32].

However, taking a closer look at the aforementioned studies, those taking non-equivalent pre-test-post-test designs (i.e. the experimental group and the control group are pre-existing groups) well outnumbered those taking counter balanced designs (e.g. [17] designed two rounds of experiments in which the group that was the experimental group in the first round became the control group in the second round). Based on this reality, this study searched and included both true experiments and quasi-experiments for meta-analysis.

In this study, two high quality databases (i.e. *Web of Science* and *Scopus*) were searched with a search formula of TITLE-ABS-KEY ("m-learning" OR mlearning OR "mobile learning" OR "mobile education" OR "m-education" OR mEducation) AND TITLE- ABS-KEY ("English learning" OR "second language learning" OR "second language acquisition" OR "foreign language learning"). The time span was set as 1900-2021, and the search
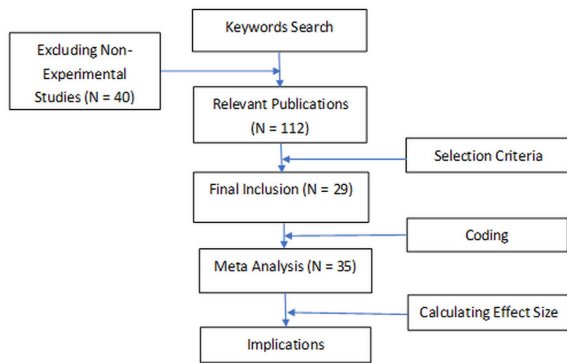
FIGURE 1. Research procedure.

yielded 152 results. An additional search was carried out in other databases such as *ERIC*, *EBSCO*, and *Google Scholar* that might store relevant articles. These results were initially screened by excluding theoretical studies, literature reviews, case studies, and pilot studies, and a total of 112 experimental and quasi-experimental studies were short-listed after duplicates were removed.

If a piece of literature was obtained from gray literature sources, it was then critically evaluated to determine the quality as gray literature might not receive the same quality appraisal as peer reviewed ones. Gray literature in this study was appraised to the same standards stipulated in *PRISMA* as those used to evaluate any other black literature. We verified the reliability of the source, the This study further set the following selection criteria:

1) The use of M-learning was the main independent variable in the study (i.e. the experimental group used mobile devices while the control group used traditional learning methods such as teacher-led, traditional, or paper-based materials). Studies were excluded when both the experimental and control groups used mobile devices for learning (e.g. a comparative study of two groups using mobile applications [34]) or a study to test learning effectiveness between genders using mobile devices [27].
2) Learning effectiveness was the main dependent variable. Studies were excluded when they investigated affective attitudes, learning attitudes, motivation interest etc. rather than learning effectiveness (e.g. [26], [35]–[38]).
3) Sufficient data were provided to calculate the effect values. Studies were excluded when they did not specify the number of participants in the experimental and control groups or the necessary information for measuring learning effectiveness, as in the cases of [39] and [40].

After applying the above criteria, 29 papers were selected for further meta-analysis. These 29 studies were conducted in the following countries or regions: Taiwan (8), the People's Republic of China (7), Turkey (6), Hong Kong (2), Iran (2), the Czech Republic (1), South Korea (1), the Netherlands (1),

and the United States (1). Fig 1 summarized our procedure of carrying out the meta-analysis.

B. MODERATORS AND CODING

1) THE SELECTION OF MODERATORS

According to [29] and [30], a number of factors (referred to as moderators) in meta-analysis can influence the independent variable and consequently predicate the final outcome of effectiveness. In our study, factors such as the subjects' grades, academic courses, teaching methods etc. were considered as moderators which might impact on the independent variable(i.e., the use of M-learning in this study). Although M-learning was noted to enhance students' English learning effectiveness in many studies, some researchers pointed out that as a tool of learning or teaching, M-learning did not guarantee increased learning effectiveness; the learning effectiveness was subject to numerous factors (variables) such as motivation, initiative, satisfaction, teacher's motivation, etc. [3], [41]. Although some aspects emerged as possible moderators influencing the effect of M-learning, such as learner satisfaction, attitude, appreciation, and motivation [28], [42]–[44], they were not included in this study as the number of such studies was insufficient to back a complete meta-analysis.

This study focused on coding 5 moderator variables, namely: learner level, mobile device, targeted areas, intervention time, and site. These 5 moderators have been noted to have effect on learners' learning outcomes in the 29 studies that were included, and they received varying degrees of attention from the authors of these studies.

In all 29 studies, the students' level of English proficiency was considered as an important factor in the study design and an important moderator when interpreting learning outcomes. For example, the researcher had to make some adjustments to the study materials because the participants were fifth grade students with insufficient English proficiency [45]. As [46] pointed out, how learners of different ages and levels were affected by the M-learning process in English as a Foreign Language (EFL) needs to receive as much attention in any such studies.

The impact of different mobile devices and applications on the effectiveness of M-learning has also gained a lot of attention from researchers [3]. In our study, mobile devices were referred to as both hardware and software devices. It was found that different types of mobile devices had different effects on learning English [27]. Similarly, it was also found that different M-learning applications led to differences in learning effectiveness mainly because of their effect on students' attention [47]. Although mobile devices were identified as having an impact on M-learning effectiveness, 87% of studies were focused on cell phones, and relatively little research has been done on other mobile devices [3]. The inclusion of this moderator in our study might provide some insight much in need.

In terms of EFL learning content, the majority of studies of English teaching and learning have been focused on 4 areas: vocabulary, speaking and listening, literacy, and general instruction [3]. Among these 4 areas, the effectiveness of M-learning on vocabulary learning received most attention from the 29 studies included [27], [48]–[52]. All of these studies basically suggested that students could make significant improvement on English vocabulary learning after using mobile devices. The inclusion of this moderator might provide further evidence for these claims.

Another variable that had an important moderating effect on the effectiveness of English learning was identified as the duration of intervention, or intervention time; and, the longer the intervention time, the more effective was the English learning [51]. However, inaccurate findings might result if the total intervention time of an experiment was not long enough [50], and the varying duration of each intervention also caused a difference in results [4]. In addition, the duration of intervention itself was influenced by different factors. For example, in one study [51], students used the M-learning device much less frequently near the end of the two-week intervention period than they did at the beginning. Obviously, the intervention time was a variable that needed further analysis.

The environment (sites) in which mobile learning is implemented also has a large impact on learning effectiveness. Among our 29 studies, some found that M-learning in classroom and off-campus real-life scenarios had a very important role in enhancing learning effectiveness [17], [46], [53], [54]. However, some found that M-learning should not receive more importance than classroom instruction [49], [55], [56]. Likewise, learning effectiveness also depended largely on students' perception of the interactive learning environment [16]. In a review of M-learning in second language acquisition, [3] found that while site was an important potential factor, the majority of studies (70%) only differentiated classroom and non-classroom learning environments, with only 30% looking at the possibility of outdoor M-learning studies. Clearly, further analysis of this study using the learning/teaching environment (site) in which M-learning was implemented as a moderator would shed new light on this issue.

2) CODING OF MODERATORS

The specific coding rules of this study on moderators are as follows:

1) *Learner Level* referred to the learning stage of the subjects, including subgroups of *junior high*, *high school*, *university*, and *elementary school*.
2) *Device* referred to both M-learning devices and the application installed on them including subgroups of *handheld devices*, *smartphone*, *iPad*, *tablet*, *PDA* (Personal Digital Assistant), and *SMS*(Short Message Services).
3) *Intervention* referred to the time interval from the beginning to the end of an experiment, including the subgroups of $1 - 4$ *weeks*, $1 - 6$ *months*, and *unspecified*.
4) *Target Areas* referred to English learning contents involved in the study, including subgroups of *reading*, *reading & listening*, *reading & writing*, *speaking*, *speaking & vocabulary*, *vocabulary*, *writing*, and *overall*.
5) *Site* referred to the environment where M-learning took place, including subgroups of *classroom*, *outside − class*, and *mixed* (two or more learning environments).

C. CODING OF EXPERIMENTS

The 29 studies were coded according to rules set aforementioned and shown in Table 1. If multiple experiments on M-learning effectiveness were reported in one study, the effect size of each experiment was entered as separate samples. For example, [17] reported the effectiveness of using tablets on English writing and English speaking respectively, so two effect sizes of the same study were recorded, one for writing and the other for speaking. As a result, 35 records were entered in total.

Two researchers then analyzed the research articles before coding them separately. The inter-coder reliability showed a high degree of agreement between the two researchers with a Cohen's kappa [57] of 94%. They then discussed the differences and reached a final consensus.

III. RESULTS

A. EFFECT SIZE AND HOMOGENEITY TEST

The results of the meta-analysis and the homogeneity test are shown in Tables 2-3. The total effect size of this study is 0.893 with a significant Cochran's Q value ($Q = 220.090$, $p < 0.001$, $df = 34$, $I^2 = 84.552$), indicating that the studies included in this meta-analysis have heterogeneous effect in the population. In other words, the variance among the studies is less likely due to sampling errors [29], [58].

I^2 values of 25%, 50%, and 75% can be considered as the division of low, medium, and high levels of heterogeneity, respectively [58]. The I^2 (*see* Table 3) of this study is 84.552, which indicates that the true variance of the effect size accounts for 84.552% of the total variance (greater than 75%) and the degree of heterogeneity of the effect size is high. This supports our choice of the random effect model. In addition, the high heterogeneity of the effect size also implies that there may be potential moderating variables for the effect of mobile English learning outcomes. Therefore, further tests for the effect of moderators, i.e., Q-tests for heterogeneity between subgroups, are needed to examine whether there are differences in effect size between subgroups.

B. PUBLICATION BIAS

The publication bias of the sample data in this study is calculated using logarithm odds ratio and Fail-safe N test [29].

This meta-analysis incorporates data from 35 studies and yields a z-value of 16.39604 and a corresponding 2-tailed p-value of 0.00000 (*see* Table 4). The Fail-safe value

TABLE 1. Coding of moderators.

Study Name	Level	Device	Intervention	Trgt areas	Site
wang2016[ba]	university	smartphone app	1-6 months	reading	class
wang2016[im]	university	smartphone app	1-6 months	reading	class
wang2016[tm]	university	smartphone app	1-6 months	reading	class
yang2015	university	handheld device	1-6 months	reading &listening	class
yao2017	university	smartphone app	1-6 months	reading &writing	class
ozer2018	university	smartphone app	1-6 months	overall	class
klimova2019	university	smartphone app	1-6 months	vocabulary	mixed
jiang2014[cb]	university	sms	1-4 weeks	vocabulary	outside-class
jiang2014 [self]	university	sms	1-4 weeks	vocabulary	outside-class
jeong2011	university	smartphone app	unspecified	overall	mixed
hayati2013 [cb]	university	sms	1-4 weeks	vocabulary	mixed
hayati2013 [self]	university	sms	1-4 weeks	vocabulary	mixed
chen&chung 2008	university	pda	1-4 weeks	vocabulary	mixed
lai2016	junior high	smartphone app	1-6 months	vocabulary	mixed
chou2017	junior high	smartphone app	1-4 weeks	overall	class
shadiev2015 [1]	junior high	tablet	1-6 months	speaking	mixed
shadiev2015 [2]	junior high	tablet	1-6 months	writing	mixed
tsai2017	high school	smartphone app	1-6 months	overall	mixed
lu2008[p1]	high school	sms	1-6 months	vocabulary	mixed
lu2008[p2]	high school	sms	1-4 weeks	vocabulary	mixed
chen2010	high school	smartphone app	1-4 weeks	vocabulary	mixed
sandberg2011 [zoo]	elementary	smartphone	1-4 weeks	speaking &vocabu-lary	mixed
sandberg2011 [zoo&home]	elementary	smartphone	1-4 weeks	speaking &vocabu-lary	mixed
zhang2019 (read)	elementary	smartphone app	1-6 months	speaking	mixed
zhang2019 [story]	elementary	smartphone app	1-6 months	speaking	mixed
zhang2019 [length]	elementary	smartphone app	1-6 months	writing	mixed
zhang2019 [pattern]	elementary	smartphone app	1-6 months	writing	mixed
hwang2013	elementary	handheld device	unspecified	vocabulary	class
reeves2017	elementary	iPad	1-6 months	speaking	class
chen2019[acq]	university	smartphone	1-4 weeks	vocabulary	mixed
chen2019[ret]	university	smartphone	1-6 months	vocabulary	mixed
demir2018	university	tablet	1-6 months	vocabulary	mixed
hsieh2015	university	tablet	1-4 weeks	vocabulary	mixed

TABLE 2. Effect size.

Model	Number Studies	Point estimate	Standard error	Variance	Lower limit	Upper limit
		Effect size and 95% confidence interval				
Fixed	35	0.784	0.050	0.002	0.687	0.882
Random	35	0.893	0.130	0.017	0.638	1.147

TABLE 3. Heterogeneity test.

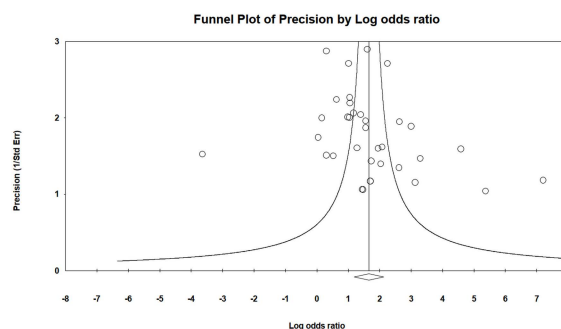| Test of null (2-Tail)		Heterogeneity			
Z-value	P-value	Q-value	df (Q)	P-value	I-squared
15.782	0.000	220.090	34	0.000	84.552
6.882	0.000				

FIGURE 2. Funnel plot of standard error.

Figure 2 shows the funnel plot of test publication bias generated by CMA3.0. As shown in the plot, the majority of studies appear toward the top of the graph and tend to cluster near the mean effect size. This indicates the absence of publication bias with the studies being distributed symmetrically about the combined effect size.

C. OVERALL EFFECT SIZE ANALYSIS

The overall effect size of the experiments included in this study is 0.893 with a Q value of $220.090(p < 0.001, df = 33)$ using the random effects model (*see* Tables 2-3). Cohen [31] suggests that the effect is small, moderate, or large when the effect size is around 0.2, 0.5, or 0.8 respectively. In this study, the effect of M-learning is estimated using 35 sample effect sizes, and the overall effect size is greater than 0.8, indicating that M-learning has a positive impact on the effectiveness of English learning. Also, the lower and upper limits of the confidence interval of this meta-analysis are 1.147 and 0.638, respectively (*see* Table 2), and the Z-value (two tails) of the Test of Null is $6.882(p < 0.001)$ (*see* Table 3). These indicate that the impact of M-learning on

(Fail-safe N) is 2415, which means that we would need to locate and include 2415 'null' studies in order for the combined 2-tailed p-value to exceed 0.050. Put another way, an average of 69.0 missing studies would be needed for every observed study in order for the effect to be nullified. Therefore, the results of Fail-safe N analysis yielded from our data suggest that the current meta-analysis is reliable and sound without publication bias.

TABLE 4. Fail-Safe N analysis.

Classic fail-safe N

Z-value for observed studies	16.39604
P-value for observed studies	0.00000
Alpha	0.05000
Tails	2.00000
Z for alpha	1.95996
Number of observed studies	35.00000
Number of missing studies that would bring p-value to > alpha	2415.00000

the English learning effectiveness is significantly higher than that of non-mobile learning. Therefore, it can be concluded that M-learning has a positive impact on the effectiveness of English learning, and its impact is higher than that of non-mobile learning.

D. ANALYSIS OF THE MODERATING EFFECTS

Table 5 shows the effect on mobile English learning outcomes brought about by the moderators, namely, *Learner Level*, *Device*, *Intervention time*, *Targeted Areas*, and *Site*. Since the CMA program understands that all the moderators included in this study are categorical variables, it automatically creates dummy codes for each category of the moderators after choosing one of them as the reference category. The groups chosen as reference groups were: *elementary school*, *SMS*, *unspecified*, *overall*, and *class* for the moderators of *Learner Level*, *Device*, *Intervention*, *Targeted Areas*, and *Site* respectively; therefore, the reference group chosen for each moderator variable does not appear in Table 5. The CMA program puts a bracket around the dummy coded variables of the same moderator, indicating that they together form a set to represent the moderator. The q and p values to the set indicated whether the dummy variables explained any substantial amount of the variance.

1) LEARNER LEVEL

In the inter-subgroup heterogeneity test of this categorical moderator, a significance Q value of $10.76 (df = 3, p < 0.05)$ has been reached, indicating that the effect of M-learning on learner achievement is significantly different among learners across all the levels, i.e. *elementary school*, *junior high*, *high school*, and *university*. Since the coefficients for *high school* and *junior high* are -0.8484 and -1.6959 respectively, the effects of M-learning on these two levels of learners are negative, which means that M-learning hinders learning outcome. For learners of the university level, the coefficient is 0.0191, indicating M-learning's positive effect in facilitating English learning. However, all the 2-sided p values for *highschool*, *junior high*, and *university* are greater than $0.05(0.5052, 0.1017, and 0.9883, respectively)$, which suggests that although the effect of M-learning is different across these categories of learners, none of the individual effects is deemed significant.

TABLE 5. Effects of moderators.

Main results for Model 1, Random effects (MM), Z-Distribution, Hedges's g

Set	Covariate	Coefficient	Standard Error	95% Lower	95% Upper	Z-value	2-sided P-value	
	Intercept	-0.9439	1.5231	-3.9291	2.0414	-0.62	0.5355	
	Learner Level: high school	-0.8484	1.2733	-3.3441	1.6473	-0.67	0.5052	
Learner Level	Learner Level: junior high	-1.6959	1.0363	-3.7270	0.3352	-1.64	0.1017	Q=10.76, df=3, p=0.0131
	Learner Level: university	0.0191	1.3015	-2.5319	2.5700	0.01	0.9883	
	Device: handheld device	3.8324	1.9489	0.0126	7.6521	1.97	0.0492	
	Device: ipad	1.9330	1.3792	-0.7702	4.6362	1.40	0.1611	
Device	Device: pda	-1.1772	0.8485	-2.8402	0.4859	-1.39	0.1653	Q=10.93, df=6, p=0.0905
	Device: smartphone	1.0018	1.4611	-1.8619	3.8655	0.69	0.4929	
	Device: smartphone app	-0.5332	0.5028	-1.5186	0.4522	-1.06	0.2889	
	Device: tablet	0.6035	1.1356	-1.6223	2.8293	0.53	0.5951	
intervention	intervention: 1-4 weeks	1.4177	1.1394	-0.8154	3.6508	1.24	0.2134	Q=2.22, df=2, p=0.3298
	intervention: 1-6 months	1.5540	1.0508	-0.5055	3.6135	1.48	0.1392	
	Targeted Area: reading	0.7440	1.0984	-1.4088	2.8968	0.68	0.4982	
	Targeted Area: reading-lis.	-3.5927	2.4759	-8.4453	1.2600	-1.45	0.1468	
	Targeted Area: reading-wri.	0.4593	1.2268	-1.9452	2.8637	0.37	0.7081	
Targeted Area	Targeted Area: speaking	-0.8659	1.0614	-2.9461	1.2143	-0.82	0.4146	Q=12.73, df=7, p=0.0791
	Targeted Area: speaking-voc.	-2.6353	2.6533	-7.8358	2.5651	-0.99	0.3206	
	Targeted Area: vocabulary	-0.3930	0.6570	-1.6806	0.8947	-0.60	0.5498	
	Targeted Area: writing	0.1841	1.2067	-2.1810	2.5492	0.15	0.8787	
Site	Site: mixed	1.6213	0.8672	-0.0785	3.3210	1.87	0.0616	Q=3.50, df=2, p=0.1741
	Site: outside-class	1.3451	1.0175	-0.6493	3.3394	1.32	0.1862	

2) DEVICES

The inter-subgroup heterogeneity test of this moderator results in a significance level of $Q = 10.93 (df = 6, p = 0.0905 > 0.05)$, indicating that the effects of using any mobile devices has no significant impact on the English learning outcome. The coefficient values of *handheld devices*, *iPads*, *smartphones*, and *tablets* are 3.8324, 1.9330, 1.0018 and 0.6035, respectively, indicating their positive effects on the learning outcome. The coefficient values of *PDAs* and *smartphone apps* are -1.1772 and -05332, respectively, indicating that their negative impact on the learning outcome. Among all these M-learning devices examined only *handheld device* (2-sided $p = 0.0492 < 0.05$) has a significant impact on the learning outcome, and the rest, i.e. *iPads* (2-sided $p = 0.1611 > 0.05$), *PDAs* (2-sided $p = 0.1653 < 0.05$), *smartphones* (2-sided $p = 0.4929 > 0.05$), *smartphone apps* (2-sided $p = 0.2889 > 0.05$) and *tablets* (2-sided $p = 0.5951 > 0.05$), no matter their impact being positive or negative.

3) INTERVENTION

The inter-subgroup heterogeneity test of this moderator reaches a Q value of $2.22 (df = 2, p = 0.3289 > 0.05)$. This indicates that the time span of implementation of M-learning has no significant effect on learning effectiveness. The coefficient values of the categories of $1 - 4$ *weeks* and $1 - 6$ *months* are 1.4177 (2-sided $p = 0.2134 > 0.05$) and 1.5540 (2-sided $p = 0.1392 > 0.05$), indicating that both intervention periods have had some positive impact on English learning outcome but neither impact is significant. The result also shows that the length of intervention time tends to be linearly related with the effect sizes, i.e., the effectiveness of M-learning increases with time being spent in M-learning. Students using mobile devices for a longer time, i.e. $1 - 6$ *months* (coefficient

value $= 1.554$), demonstrated better learning outcomes than those using mobile devices for a short term, i.e. $1 - 4$ *weeks* (coefficient value $= 1.4177$).

4) TARGETED AREAS

The inter-subgroup heterogeneity test of this moderator shows a significant Q value of $12.73(df = 7, p = 0.0791 > 0.05)$, indicating that M-learning's effect has no significant impact on English learning outcome across the targeted learning areas. However, M-learning has shown some positive effects on English *reading*, *reading & writing*, and *writing* with respective coefficient values of 0.7440, 0.4593, and 0.1840. M-learning has shown some negative effect on *reading & listening*, *speaking*, *speaking & vocabulary*, and *vocabulary*, with respective coefficient values of -3.5927, -0.8659, -2.6353, and -0.3930. However, none of these impacts, being positive or negative, is statistically significant from others where all the 2-sided p values are greater than 0.05.

5) SITE

The inter-subgroup heterogeneity test of this moderator shows $Q = 3.50(df = 2, p = 0.1741 > 0.05)$, indicating that the site where M-learning is implemented has no significant impact on learning effectiveness. The coefficient values for the categories of *mixed* and *outside − class* are 1.6213 (2-sided $p = 0.0616 > 0.05$) and 1.3451 (2-sided $p = 0.1862 > 0.05$), respectively, indicating that there is no significant difference between implementing M-learning outside the class or in a mixed way, and each approach would help with learning effectiveness.

IV. DISCUSSIONS AND LIMITATIONS

Applying the random effects model to 35 samples in meta-analysis, this study has come to a conclusion on the positive effectiveness of mobile language learning with an effect size of 0.893. The results thus lend support to the proposition that M-learning is more effective than traditional learning and that M-learning can be used as a teaching method to facilitate English learning to occur. A large number of previous studies of both experimental and non-experimental nature reported positive effects of M-learning on English learning [4], [8], [11], [14], [18]. However, some studies found that M-learning was not effective or even had negative effects on English learning [21], [22]. The effect size of mobile English learning was to some extent calculated in a few meta-analyses [6], [12]. In a study of M-learning of various academic courses including English, an effect size of 0.482 of M-learning was reported, but it did not further provide the effect size of individual courses [6]. Another study [12] classified the effect sizes of 59 samples into three categories: large, small, and negative, but did not report a synthesized effect size (i.e., 10 studies in her analysis had effect sizes greater than 1.0, and more studies had effect sizes less than 0.5). Our study is the first to report on the effect size of mobile English learning, and the finding of an effect size of 0.893 strongly supports the idea that M-learning enhances

learning effectiveness. Although many studies in second language research are of quasi-experiments, which may weaken the strength of the evidence, our study has drawn conclusions through a carefully conducted meta-analysis, which includes a synthesis of effect sizes from each study, a heterogeneity test, and a publication bias test following *PRISMA* guidelines [33]. In fact it is the advantage of meta-analysis over other systematic review methods used to synthesize the effect sizes of different studies, and this study has made a meaningful attempt in the examination of the effect size of mobile English learning.

This study has also specifically investigated the moderating effects of 5 moderators on the mobile English learning outcome, including learner levels, devices, intervention time, targeted learning areas, and implementation sites. In terms of learner levels, M-learning has a significant positive effect on university learners and a negative effect on the junior and high student cohorts. Previous studies have pointed out the disproportionately greater number of studies on college and adult learners than on language learners at other levels [3], [13], [24], but our study suggests that the issue of proportionality may not be true or may have been improved in recent years as quite a number of studies on high school and junior high school learners have emerged in our datasets.

Our finding of significant difference in the effect of M-learning between university students and learners of other levels is not unexpected, as college students have long been considered to be more motivated and capable of learning than students at other levels [13], [46]. In addition, differences in motivation, language background, learning style and ability, and instructional environment can also lead to differences in learning effectiveness [5], [12], [13]. Therefore, teachers, especially those teaching middle and high schools, should be more cautious when using M-learning. We believe that maintaining communication and feedback is a good suggestion [13]. In addition, teachers can pay close attention to the development of students' learning communities, help students develop the right attitude towards M-learning, and update learning content in real time, which are all key factors to the eventual or continued success of M-learning [5].

In terms of devices, this study shows that there is significant difference in the effects of hardware devices on the effectiveness of M-learning, with *handheld devices*, *iPads*, *smartphones* and *tablets* having positive effects on learning effectiveness, and *PDAs* and *smartphone apps* having negative effects. Since definitions of devices are inconsistent across studies, i.e. iPads in fact belong to tablets, this study has followed the definitions and names used by the included studies without attempting to provide a unified definition, which may have resulted in some inconsistencies in the findings. Nonetheless, the finding of negative effect of the smartphone app suggests that attempts to integrate multiple technologies to develop new mobile applications into the overall language education process may have a very limited and possibly counter-productive impact on possible student learning outcomes. Although a number of studies

involving the development of products for language teaching and learning have reported that new apps can promote positive learning outcomes (e.g. [52], [56], [59]), the results of our meta-analysis do not support these claims. This is perhaps due to these new apps rarely becoming commercially available software or tools, making it impossible to compare them with other products. The other possible reason is that the researchers themselves are the developers in those studies, which may have an impact on their results, as pointed out by [13] and [6]. On the other hand, our findings show that generic products can also facilitate M-learning by relying on sound design. Therefore, pedagogical approaches to effective M-learning may incorporate the characteristics of mobile devices, including accessibility, affordability, and the aforementioned student initiative, authenticity of the learning environment [13], [25], interactivity of the learning process, and contextualization of teaching activities [5], [6].

Many previous studies have indicated that the length of time spent learning with mobile devices (intervention time) is not associated with better English learning outcomes [4], [50], [51]. In one specific example [51], the researchers noticed that 5th grade students used M-learning devices far less frequently after 2 weeks than when they first started. This study shows that the effect of using mobile devices for $1 - 6$ months is greater than that of for 1-4 weeks. The findings of our study suggest that the time span over which M-learning is implemented has no significant difference in learning outcomes, but the time span is linearly related to learning outcomes – a longer span leads to a better effect. This result basically supports findings of the previous studies. However, as many studies included in this study did not report specific intervention times (coded as unspecific), the effects of intervention time on mobile English learning could not be accurately evaluated. Although this issue has received some attention, researchers have called for more studies to look at the long-term effect of M-learning [24], [25]. The results of our study have, to some extent, revealed the moderating effect of intervention time in M-learning, and it is hoped that in future research more attention is paid to this variable in true experiments. In practical language teaching, the results of this study suggest that M-learning is best achieved when it is designed in a timely manner, taking into account the specific pace of instruction, rather than requiring students to use mobile devices on a permanent basis. Teachers may consider combining M-learning with self-directed learning, task-driven learning, cooperative learning, and blended learning, in which certain positive learning outcomes may be achieved in as short as 4 weeks.

This study has also found that the effects of using M-learning on English *reading* and English *reading & writing* are positive, but they have turned negative in English *reading & listening*, *speaking*, *speaking & vocabulary*, and *vocabulary*. In specific, it appears that M-learning can help students improve their reading comprehension scores in a variety of English tests or exams. For example, Wang [59] has found that students using M-learning have achieved better reading comprehension scores on multiple-choice tests in TOEIC because they could remember the spelling of English words more accurately than students in the control group. Yao [52] has found that students in the M-learning experimental group have used more English idioms than those in the control group, indicating that the former have obtained a deeper memory for the idioms. In [36], students in the M-learning group have been found to be able to use a greater variety of English vocabulary than students doing traditional and non-mobile learning. The aforementioned studies have shown the effects of M-learning on specific language phenomena. Because much of the previous research has focused more on teaching and learning of general language areas than on the effects of each specific learning area [24], this study is the first to provide such information. According to some previous studies [5], [6], [12], there is a complex set of factors that influence the effectiveness of M-learning, and pedagogical goals and content are only parts of them. It is easier for learners to make progress in English reading and writing with M-learning than in listening and speaking as the latter requires building a learner community in which learners can maintain practice [5]. In addition, our result may be due to the fact that reading and writing are easier to evaluate than listening and speaking; in terms of the technical skills required, all that is needed to handle mobile reading and writing is a simple document editing application, a reasonable network connection, and a good access to the Internet. Clearly, more research is needed in this area.

The results also surprised us by showing a negative effect of M-learning in terms of *speaking & vocabulary* and *vocabulary*. In previous studies, the effectiveness of M-learning on vocabulary learning has received the most attention, and many studies reporting the development of M-learning tools have claimed that students made significant improvements in their English vocabulary after using tools [3], [24]. Since M-learning of vocabulary is usually the first to be completed and reported in the development of more advanced M-learning systems, many such studies only serve as a pilot study in which neither true nor quasi experiments were carried out [8], [28], [60], [61]. As a result of not using experiments, they have been excluded from this study. When the results of other experimental studies are taken into account, M-learning of vocabulary has been found not as effective as claimed by the non-experimental studies. This may offer a partial explanation of the differences between our studies and the previous ones. Another possible explanation is that many of the experimental studies included in this study are aimed at university students, for whom vocabulary learning is not their main English learning purpose, resulting in ineffective vocabulary learning outcome. From the perspective of this meta-analysis, the results of this study again show the power of meta-analysis to integrate different studies and provide new insights into the effectiveness of mobile English learning for practical purposes.

In terms of implementation settings, our results have shown that the adoption of M-learning in a *mixed* or

outside − class learning environment has no significant impact on learning outcomes, with the degree of facilitation being more pronounced in the mixed learning environment. Previous studies do not have a consistent view on the context in which M-learning is implemented, and the majority of them only classify M-learning as classroom and non-classroom status [3]. This result of our study has implications for the development of M-learning applications, i.e., mixed learning environments should be a priority when developing M-learning applications. It is reassuring that many recent M-learning applications are being developed based on mixed learning environments [16], [56].

Finally, our results show that the effectiveness of mobile English learning is influenced by a variety of moderators, and some of the moderators even have ineffective or negative impact on the learning outcome. However, when a moderator is found to have ineffective or negative impact, it does not mean that it is a failing variable. For example, our research shows that M-learning within 1-6 months is more effective than that within 1-4 weeks, but this does not necessarily mean that short-term M-learning is a bad strategy. In fact, some language areas may improve in the short term. In the long run, moderators showing ineffective or negative impact on learning outcome may translate into positive impact later on, which is a view supported by many scholars who have proposed longitudinal studies of learning outcomes [5], [24].

In summary, this study is based on a meta-analysis of 29 experimental studies of M-learning in the field of EFL. Although the results are satisfactory and shed some light into the effectiveness of M-learning, there are still some limitations, as follows. The sample number for analyzing the effect sizes of subgroups is relatively small, particularly for those moderators with more categories (i.e., the moderators of device and targeted areas). The solution is to include a greater number of true or quasi experiments, which also carries the risk of getting the data contaminated by unreliable studies [29]. Second, although this study has taken into account 5 moderators, many more have been left out such as learner attitude, gender, socio-economic status, income, and maturity as researchers do not always explicitly report their variables [62]. Furthermore, the effect of language learning usually shows in the long run. The intervention time of experiments included in this study ranges from a few weeks to 6 months, and thus long-term effects of M-learning, together with the effects of the 5 moderators, have not been established. As a result of these stated limitations, the relevant results should be interpreted with caution.

V. CONCLUSION

This study has examined the overall effect of mobile English learning and the specific effects of moderating factors, namely, learners' levels, M-learning devices, intervention time, targeted English learning areas, and the site where M-learning takes place. Although further research is needed, the evidence from this study clearly supports that using M-learning has a positive effect on English teaching and learning. It is hoped that the findings of this study may provide useful information for language teachers and learners of all levels who are considering going mobile in teaching or learning. The results may also help researchers interested in exploring mobile technology in the educational context to gain a bird's eye view of the current research development in the EFL field.

REFERENCES

[1] W.-H. Wu, Y.-C. J. Wu, C.-Y. Chen, H.-Y. Kao, C.-H. Lin, and S.-H. Huang, "Review of trends from mobile learning studies: A meta-analysis," *Comput. Educ.*, vol. 59, no. 2, pp. 817–827, Sep. 2012.

[2] Q. Wu, "Designing a smartphone app to teach English (L2) vocabulary," *Comput. Educ.*, vol. 85, pp. 170–179, Jul. 2015.

[3] V. Persson and J. Nouri, "A systematic review of second language learning with mobile technologies," *Int. J. Emerg. Technol. Learn.*, vol. 13, no. 2, pp. 188–210, 2018.

[4] P.-N. Chou, C.-C. Chang, and C.-H. Lin, "BYOD or not: A comparison of two assessment strategies for Student learning," *Comput. Hum. Behav.*, vol. 74, pp. 63–71, Sep. 2017.

[5] M. Alrasheedi and L. F. Capretz, "Determination of critical success factors affecting mobile learning: A meta-analysis approach," *Turkish Online J. Educ. Technol.*, vol. 14, no. 2, pp. 41–51, 2015.

[6] S. Tingir, B. Cavlazoglu, O. Caliskan, O. Koklu, and S. Intepe-Tingir, "Effects of mobile devices on K-12 students' achievement: A meta-analysis: Effects of mobile devices," *J. Comput. Assist. Learn.*, vol. 33, no. 4, pp. 355–369, Aug. 2017.

[7] J. L. Reeves, G. A. Gunter, and C. Lacey, "Mobile learning in pre-kindergarten: Using student feedback to inform practice," *Educ. Technol. Soc.*, vol. 20, no. 1, pp. 37–44, 2017.

[8] R. Shadiev, W.-Y. Hwang, Y.-M. Huang, and T.-Y. Liu, "Facilitating application of language skills in authentic environments with a mobile learning system," *J. Comput. Assist. Learn.*, vol. 34, no. 1, pp. 42–52, Feb. 2018.

[9] R. Shadiev, W.-Y. Hwang, Y.-M. Huang, and T.-Y. Liu, "Cognitive diffusion model: Facilitating EFL learning in an authentic environment," *IEEE Trans. Learn. Technol.*, vol. 10, no. 2, pp. 168–181, Apr. 2017.

[10] Z. Sun, C.-H. Lin, J. You, H. J. Shen, S. Qi, and L. Luo, "Improving the English-speaking skills of young learners through mobile social networking," *Comput. Assist. Lang. Learn.*, vol. 30, nos. 3–4, pp. 304–324, May 2017.

[11] J. Traxler, E. Barcena, and J. G. Laborda, "Mobile technology for foreign language teaching: Building bridges between non-formal and formal scenarios," *J. Universal Comput. Sci.*, vol. 21, no. 10, pp. 1234–1247, 2015.

[12] M. M. Malissa, "Technology and language – what works and what does not: A meta-analysis of blended learning research," *J. Asia TEFL*, vol. 15, no. 2, pp. 365–382, Jun. 2018.

[13] Y. Zhao, "Recent developments in technology and language learning: A literature review and meta-analysis," *CALICO J.*, vol. 21, no. 1, pp. 7–27, Jan. 2013.

[14] P.-L. Liu and C.-J. Chen, "Learning English through actions: A study of mobile-assisted language learning," *Interact. Learn. Environ.*, vol. 23, no. 2, pp. 158–171, Mar. 2015.

[15] M. Saran, G. Seferoglu, and K. Cagiltay, "Mobile assisted language learning: English pronunciation at learners' fingertips," *Eurasian J. Educ. Res.*, vol. 34, pp. 97–114, 2009.

[16] A. Lai, "Mobile immersion: An experiment using mobile instant messenger to support second-language learning," *Interact. Learn. Environ.*, vol. 24, no. 2, pp. 277–290, Feb. 2016.

[17] R. Shadiev, W.-Y. Hwang, Y.-M. Huang, and T.-Y. Liu, "The impact of supported and annotated mobile learning on achievement and cognitive load," *J. Educ. Technol. Soc.*, vol. 18, no. 4, pp. 53–69, 2015.

[18] S. F. A. Hossain, M. Nurunnabi, K. Hussain, and S. K. Saha, "Effects of variety-seeking intention by mobile phone usage on university students' academic performance," *Cogent Educ.*, vol. 6, no. 1, Jan. 2019, Art. no. 1574692.

[19] S. F. A. Hossain, Y. Ying, and S. K. Saha, "Systematic mobile device usage behavior and successful implementation of tpack based on university students need," in *Advances in Computer Vision* (Advances in Intelligent Systems and Computing), vol. 1, K. Arai and S. Kapoor, Eds. Cham, Switzerland: Springer, 2020, pp. 729–746.

[20] E. B. Başoğlu and O. Akdemir, "A comparison of undergraduate students' English vocabulary learning: Using mobile phones and flash cards," *Turkish Online J. Educ. Technol.*, vol. 9, no. 3, pp. 1–7, 2010.

[21] D. Cojocnean, "Factors determining Students' low usage of mobile tools in their English vocabulary learning," *Porta Linguarum Revista Interuniversitaria de Didáctica de las Lenguas Extranjeras*, vol. Monográfico I, pp. 31–43, Sep. 2016.

[22] R. Dashtestani, "Efl teachers' and students' perspectives on the use of electronic dictionaries for learning English," *CALL-EJ*, vol. 14, no. 2, pp. 51–65, 2013.

[23] P. C. Pollara, "Mobile learning in higher education: A glimpse and a comparison of student and faculty readiness, attitudes and perceptions," Ph.D. dissertation, Dept. Educ. Theory, Policy, Pract., Louisiana State Univ., Baton Rouge, LA, USA, 2011.

[24] M. M. Elaish, L. Shuib, N. A. Ghani, E. Yadegaridehkordi, and M. Alaa, "Mobile learning for English language acquisition: Taxonomy, challenges, and recommendations," *IEEE Access*, vol. 5, pp. 19033–19047, 2017.

[25] R. Shadiev, W.-Y. Hwang, and Y.-M. Huang, "Review of research on mobile language learning in authentic environments," *Comput. Assist. Lang. Learn.*, vol. 30, nos. 3–4, pp. 284–303, May 2017.

[26] L. K. Cheng, A. Selamat, F. Puteh, and F. Mohamed, "A review of recent methodologies, technologies and usability in English language content delivery," *Jurnal Teknologi*, vol. 78, nos. 12–3, pp. 1–11, Dec. 2016.

[27] Z. Yu, "A systematic review on mobile technology-assisted English learning," *Int. J. e-Collaboration*, vol. 15, no. 4, pp. 71–88, Oct. 2019.

[28] H.-C. Hung and S. S.-C. Young, "An investigation of game-embedded handheld devices to enhance English learning," *J. Educ. Comput. Res.*, vol. 52, no. 4, pp. 548–567, Jul. 2015.

[29] M. Borenstein, L. V. Hedges, J. P. T. Higgins, and H. R. Rothstein, *Introduction to Meta-Analysis*. Chichester, U.K.: Wiley, 2009.

[30] N. A. Card, *Applied Meta-Analysis for Social Science Research* (Methodology in the Social Sciences). New York, NY, USA: The Guilford Press, 2011.

[31] J. Cohen, *Statistical Power Analysis for the Behavioral Sciences*, 2nd ed. Mahwah, NJ, USA: Lawrence Erlbaum Associates, 1988.

[32] M. Borenstein, L. V. Hedges, J. P. Higgins, and H. R. Rothstein, "Comprehensive meta analysis," Biostat Inc., Englewood, NJ, USA, Tech. Rep., 2013.

[33] M. J. Page *et al.*, "The PRISMA 2020 statement: An updated guideline for reporting systematic reviews," *Systematic Rev.*, vol. 10, no. 89, pp. 1–11, 2021.

[34] H. Yang, "The effects of attention cueing on English reading on mobile phones," *Frontiers Educ. China*, vol. 13, no. 3, pp. 315–345, Sep. 2018.

[35] C.-C. Chang, C.-F. Yan, and J.-S. Tseng, "Perceived convenience in an extended technology acceptance model: Mobile technology and English learning for college students," *Australas. J. Educ. Technol.*, vol. 28, no. 5, pp. 809–826, Jul. 2012.

[36] D. Zhang and P. Pérez-Paredes, "Chinese postgraduate EFL learners' self-directed use of mobile English learning resources," *Comput. Assist. Lang. Learn.*, vol. 34, no. 8, pp. 1128–1153, Nov. 2021.

[37] O. Viberg and Å. Grönlund, "Cross-cultural analysis of users' attitudes toward the use of mobile devices in second and foreign language learning in higher education: A case from Sweden and China," *Comput. Educ.*, vol. 69, pp. 169–180, Nov. 2013.

[38] S.-H. Yang, "Exploring college students' attitudes and self-efficacy of mobile learning," *Turkish Online J. Educ. Technol.*, vol. 11, no. 4, pp. 148–154, 2012.

[39] Y. Chen, C. L. Carger, and T. J. Smith, "Mobile-assisted narrative writing practice for young English language learners from a funds of knowledge approach," *Lang. Learn. Technol.*, vol. 21, no. 1, pp. 28–41, 2017.

[40] M. Kuimova, D. Burleigh, H. Uzunboylu, and R. Bazhenov, "Positive effects of mobile learning on foreign language learning," *TEM J.*, vol. 7, no. 4, pp. 837–841, 2018.

[41] F. Moreira, M. J. Ferreira, C. P. Santos, and N. Durão, "Evolution and use of mobile devices in higher education: A case study in Portuguese higher education institutions between 2009/2010 and 2014/2015," *Telematics Inform.*, vol. 34, no. 6, pp. 838–852, Sep. 2017.

[42] L. Hsu and S.-N. Lee, "Learning tourism English on mobile phones: How does it work?" *J. Hospitality Leisure Sport Tourism*, vol. 10, no. 2, pp. 85–94, Oct. 2011.

[43] A. Andujar, M. S. Salaberri-Ramiro, and M. S. C. Martínez, "Integrating flipped foreign language learning through mobile devices: Technology acceptance and flipped learning experience," *Sustainability*, vol. 12, no. 3, p. 1110, Feb. 2020.

[44] Y. Bayyurt, G. Erçetin, and N. B. Karataş, "The stages in mobile-assisted language learning material development," in *Mobile as a Mainstream—Towards Future Challenges in Mobile Learning*, vol. 479. M. Kalz, Y. Bayyurt, and M. Specht, Eds. Istanbul, Turkey: Springer, 2014, pp. 339–350.

[45] H. Zhang, W.-Y. Hwang, S.-Y. Tseng, and H. S. L. Chen, "Collaborative drama-based EFL learning in familiar contexts," *J. Educ. Comput. Res.*, vol. 57, no. 3, pp. 697–722, Jun. 2019.

[46] O. Ozer and F. Kılıç, "The effect of mobile-assisted language learning environment on EFL Students' academic achievement, cognitive load and acceptance of mobile learning tools," *EURASIA J. Math., Sci. Technol. Educ.*, vol. 14, no. 7, pp. 2915–2928, May 2018.

[47] L. E. Levine, B. M. Waite, and L. L. Bowman, "Mobile media use, multitasking and distractibility," *Int. J. Cyber Behav., Psychol. Learn.*, vol. 2, no. 3, pp. 15–29, Jul. 2012.

[48] C.-M. Chen and C.-J. Chung, "Personalized mobile English vocabulary learning system based on item response theory and learning memory cycle," *Comput. Educ.*, vol. 51, no. 2, pp. 624–645, Sep. 2008.

[49] J. Jiang, "A probe into the effectiveness of non-English majors' SMS-based English idiom acquisition in China," *Int. J. e-Collaboration*, vol. 10, no. 3, pp. 30–43, Jul. 2014.

[50] M. Lu, "Effectiveness of vocabulary learning via mobile phone," *J. Comput. Assist. Learn.*, vol. 24, no. 6, pp. 515–525, Oct. 2008.

[51] J. Sandberg, M. Maris, and K. de Geus, "Mobile English learning: An evidence-based study with fifth graders," *Comput. Educ.*, vol. 57, no. 1, pp. 1334–1347, Aug. 2011.

[52] C.-B. Yao, "Constructing a user-friendly and smart ubiquitous personalized learning environment by using a context-aware mechanism," *IEEE Trans. Learn. Technol.*, vol. 10, no. 1, pp. 104–114, Jan./Mar. 2017.

[53] C.-M. Chen and Y.-L. Li, "Personalised context-aware ubiquitous learning system for supporting effective English vocabulary learning," *Interact. Learn. Environments*, vol. 18, no. 4, pp. 341–364, Dec. 2010.

[54] W.-Y. Hwang and H. S. L. Chen, "Users' familiar situational contexts facilitate the practice of EFL in elementary schools with mobile devices," *Comput. Assist. Lang. Learn.*, vol. 26, no. 2, pp. 101–125, Apr. 2013.

[55] A. Hayati, A. Jalilifar, and A. Mashhadi, "Using short message service (SMS) to teach English idioms to EFL students," *Brit. J. Educ. Technol.*, vol. 44, no. 1, pp. 66–81, Jan. 2013.

[56] B. Klimova, "Impact of mobile learning on students' achievement results," *Educ. Sci.*, vol. 9, no. 2, pp. 1–8, 2019.

[57] L. Cohen, L. Manion, and K. Morrison, *Research Methods in Eduction*. New York, NY, USA: Routledge, 2007.

[58] J. P. T. Higgins, S. G. Thompson, J. J. Deeks, and D. G. Altman, "Measuring inconsistency in meta-analyses," *Brit. Med. J.*, vol. 327, no. 7414, pp. 557–560, 2003.

[59] Y.-H. Wang, "Integrating self-paced mobile learning into language instruction: Impact on reading comprehension and learner satisfaction," *Interact. Learn. Environ.*, vol. 25, no. 3, pp. 397–411, Apr. 2017.

[60] A. Rezaei, N. Mai, and A. Pesaranghader, "The effect of mobile applications on English vocabulary acquisition," *Jurnal Teknologi*, vol. 68, no. 2, pp. 73–83, Apr. 2014.

[61] Z. Yu, "Gender differences in cognitive loads, attitudes, and academic achievements in mobile English learning," *Int. J. Distance Educ. Technol.*, vol. 17, no. 4, pp. 21–35, Oct. 2019.

[62] A. A. Bartolucci and W. B. Hillegass, "Overview, strengths, and limitations of systematic reviews and meta-analyses," in *Evidence-Based Practice*, F. Chiappelli and X. M. C. Brant, Eds. Heidelberg, Germany: Springer, 2010, pp. 17–33.

MENG-LIN CHEN was born in Kaohsiung, Taiwan, in 1968. She received the master's degree in education from Shenandoah University, VA, USA, in 1994, and the Ph.D. degree in education from La Sierra University, CA, USA, in 2005, both focused on teaching English to speakers of other languages.

She has been teaching English and translation for more than 15 years. She is now an Associate Professor at the Department of Translation and Interpretation Studies, Chang Jung Christian University, Tainan, Taiwan. She has recently published several articles in *Babel*, *Publications*, and *Scientometrics*. Her research interests include SLA studies, corpus translation studies, and translation studies for special purposes. She is a Full Member of the Taiwan Association of Translation and Interpretation.

• • •