Title	Apps for English language learning: A systematic review
Author(s)	Fei V Lim and Weimin Toh

This is the published version of the following article:

Lim, F. V., & Toh, W. (2024). Apps for English language learning: A systematic review. *Teaching English with Technology, 24*(1), 79–98.

Lim, F. V., & Toh, W. (2024). APPS for ENGLISH LANGUAGE LEARNING: A SYSTEMATIC REVIEW. *Teaching English with Technology*, *24*(1), 79–98.
https://doi.org/10.56297/FSYB3031/GAQR3589

ISSUES

APPS for ENGLISH LANGUAGE LEARNING: A SYSTEMATIC REVIEW

Fei V Lim[1], Weimin Toh[1]

[1] National Institute of Education, Nanyang Technological University

Keywords: educational technology, digital media, systematic literature review, language learning

https://doi.org/10.56297/FSYB3031/GAQR3589

Teaching English with Technology

Vol. 24, Issue 1, 2024

This article reports on a systematic review of research studies published from 2010 to 2021 on the use of apps for learning in the secondary English as a Second Language (ESL) classroom. We extracted relevant information such as the studies' country/region, research design, sample size, students' age, apps' names to conduct a thematic analysis to associate the types of apps and their learning outcomes. The findings suggest that quiz apps support vocabulary acquisition, puzzle apps support vocabulary and grammar learning, platform apps support reading and writing development, augmented reality apps support increased engagement, and virtual reality apps support development of listening and speaking skills. The factors involved in the effective use of apps for learning in the English classroom include the use of quality apps to support the teacher's pedagogy, the profile and readiness of students to engage in digital learning, and the recognition that a classroom ecology is needed for effective integration of digital resources for teaching and learning.

1. Introduction

The increasing availability and prevalence of the use of apps for learning on digital devices in the second language (L2) English classroom today invites the question of their effectiveness. This emerging field of research has been reported in the research literature, ranging from small-scale exploratory studies (e.g., Lim & Toh, 2022; L.-H. Wong & Looi, 2010), to large-scale research projects (e.g., Cheung & Slavin, 2012; Mifsud et al., 2013). We perform a systematic review of the recent studies on apps for learning in the L2 English classroom and identify the specific types of apps and their affordances for various forms of language learning. Both the English as Second Language (ESL) and English as Foreign (EFL) classroom studies are included to reflect the focus on language learning for non-native speakers.

1.1. Research question, definition of apps, and rationale for study

Our systematic review is guided by the research question: Based on a systematic review of international studies that explicitly reference apps for learning, published in the last 12 years from 2010 to 2021, what are the typical learning gains associated with each category of apps used in the L2 English classroom?

In our review, apps are theoretically defined along a continuum from those that can foster learning through lower order thinking skills to others that can promote higher order thinking skills. We define apps for learning as a broad category of apps that can be sub-classified into (1) quiz apps, (2) puzzle apps, (3) platform apps, (4) augmented reality apps, and (5) virtual reality apps (Carrier et al., 2017; Perrotta et al., 2013). The desired learning outcomes in the L2 English classroom, while recognised to be curriculum-specific, are broadly

classified as related to English learning in the areas of vocabulary and grammar learning, as well as reading, writing, speaking, and listening skills. The rationale for our study is that there have been no up-to-date systematic reviews done on apps for learning in the English classroom. Due to rapid developments in technology, there is a need for teachers to understand the affordances of emerging technologies and how they can be used in the English classroom.

2. Literature review

Macaro et al. (2012) conducted a systematic review of computer-assisted language learning technologies in L2 teaching in the primary and secondary English classrooms. They found that technology use can qualitatively change students' learning experiences and there is evidence that learners possessed positive attitudes towards the use of technologies for language learning. However, their systematic review was conducted on studies from the 1990s to 2010 and there has not been a recent review of apps for learning in the English classroom. Gangaiamaran and Pasupathi (2017) discussed the different types of apps for learning and their appropriateness for different levels of learning. Apps useful to students in the primary L2 English classroom context are those that focus on basic language skills, such as learning the letters of the alphabet, learning the sound of words and its parts, and practising writing by tracing the letters through games and flashcards. In the secondary L2 English classroom, apps focus on grammar and compositions. Apps used in the tertiary contexts tend to be those that provide access to authentic material such as live streams, English songs, radio, and English news to develop listening skills. These apps often have affordances such as social connectivity, gamified elements, and offer the user a certain degree of personalisation (Brand & Kinash, 2010).

We complement the findings from Gangaiamaran and Pasupathi (2017) by examining more closely the types of apps used for learning in the secondary L2 English classroom. The choice to focus on the secondary L2 English classroom was influenced by the higher availability of relevant studies on ESL/EFL in the secondary school context, as well as the recognition that secondary school ESL/EFL teachers are in a better position to implement the use of apps for learning in the classroom (Vaish & Towndrow, 2010).

3. Methods

We performed a high-level review of the available, relevant information to extract and analyse data to address a specific research question (Schaefer & Myers, 2017). Our review was guided by the stages of research synthesis (Cooper, 2016): (1) research question formulation, (2) selection of journals, (3) definition of inclusion and exclusion criteria for studies, (4) definition of apps' categories for analysis, (5) specification of search terms, (6) data extraction and coding, and (7) synthesis. Following the synthesis, the review analysis of the results is reported, and discussion of the findings, trends, and conclusions regarding the "Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA)" statement (Moher et al., 2009) is conducted.

3.1. Selection of journals

Relevant papers were selected from the list of top 20 journals using the Google Scholar h5-Index for the category "Educational Technology". These journals include *Computer Assisted Language Learning*, *Educational Technology Research and Development*, *British Journal of Educational Technology*, *Interactive Learning Environments*, and *Education and Information Technologies*. However, for a systematic review, it is not sufficient for us to focus only on higher ranked academic journals as we may fail to cover important pedagogical studies on the use of apps in the English classroom if we only focus on them. Therefore, we also included other Q1 journals such as *Teaching English with Technology* and *English Language Teaching*, as well as book chapters, and grey literature such as conference papers and unpublished PhD and Master theses in the systematic review (Paez, 2017).

3.2. Inclusion and exclusion criteria

The inclusion criteria that were used to guide the article selection process included the following:

- Articles published between 2010 to 2021.
- Articles published in the English language.
- Articles reporting empirical studies published in journal papers, conference papers, book chapters, and theses.
- Peer-reviewed articles.
- Articles focusing on the use of apps for learning in the ESL/EFL classrooms.
- Learning outcomes reported include ESL/EFL learning.
- Studies describing applications of apps for learning in secondary and high schools.
- Studies describing the app's affordances that are relevant for the ESL/EFL learning outcomes.

The exclusion criteria included the following:

- Articles that are theoretical in nature.
- The reported studies are conducted in the home context.
- Non-English language articles.
- Articles that reported on studies using non-app like display devices such as interactive white boards and social networking websites.

3.3. Search strategy

The search strategy included the manuscript selection process. Google Scholar and education-related databases such as ProQuest, ERIC (Education Resources Information Center), and ScienceDirect were searched using the following search terms: "game-based curriculum", "apps for learning", and "game-based learning". These search terms were combined with AND and OR Boolean operators and the minus sign for excluding specific terms to enable us to focus the search and connect various pieces of information to find exactly what we were looking for.

The search string used for Google Scholar was "("game-based curriculum" OR "game-based learning" OR "apps for learning" OR "educational software") AND (ESL OR EFL) AND ("secondary classroom" OR "secondary student" OR "secondary school") AND student -home -science". The search string used for ScienceDirect was "("game-based curriculum" OR "game-based learning" OR "apps for learning" OR "educational software" OR quiz* OR "puzzle apps" OR "platform apps" OR "augmented reality apps" OR "virtual reality apps") AND (ESL OR EFL) AND ("secondary classroom" OR "secondary student" OR "secondary school") AND student -home -science". The search string used for ProQuest was "((learning N/4 (app? OR application?)) OR (language N/4 (app? OR application?)) OR "educational software" OR "game-based learning" OR quiz* OR "puzzle apps" OR "platform apps" OR "augmented reality apps" OR "virtual reality apps") AND (ESL OR EFL) AND ("secondary classroom" OR "secondary student" OR "secondary school") AND (language OR linguistic OR vocab* OR literacy OR read* OR spell* OR writ* OR listen* OR speak* OR speech OR talk* OR narrative* OR meta-linguistic)".

The search string used for ERIC via EBSCO databases was "((learning N(4) (app? OR application?)) OR (language N(4) (app? OR application?)) OR "educational software" OR "game-based learning" OR quiz* OR "puzzle apps" OR "platform apps" OR "augmented reality apps" OR "virtual reality apps") AND (ESL OR EFL) AND ("secondary classroom" OR "secondary student" OR "secondary school") AND (language OR linguistic OR vocab* OR literacy OR read* OR spell* OR writ* OR listen* OR speak* OR speech OR talk* OR narrative* OR meta-linguistic)". The last search was conducted on Dec 12, 2021. Two researchers read through the full texts of the papers and selected the relevant papers based on the inclusion and exclusion criteria. During the search cycle, 35 articles were gathered, and three articles (Cabrera et al., 2018; Quiroz et al., 2021; Shaikh et al., 2021) were added after peer review.

We searched for articles which were written in English and appeared in peer-reviewed scholarly journals, book chapters, conference papers, and theses between January 2010 and December 2021. The search included 12 years of studies for two reasons. First, increased digitalisation (Gustafsson, 2021; Olofsson & Lindberg, 2021) has led to the adoption of new technologies into

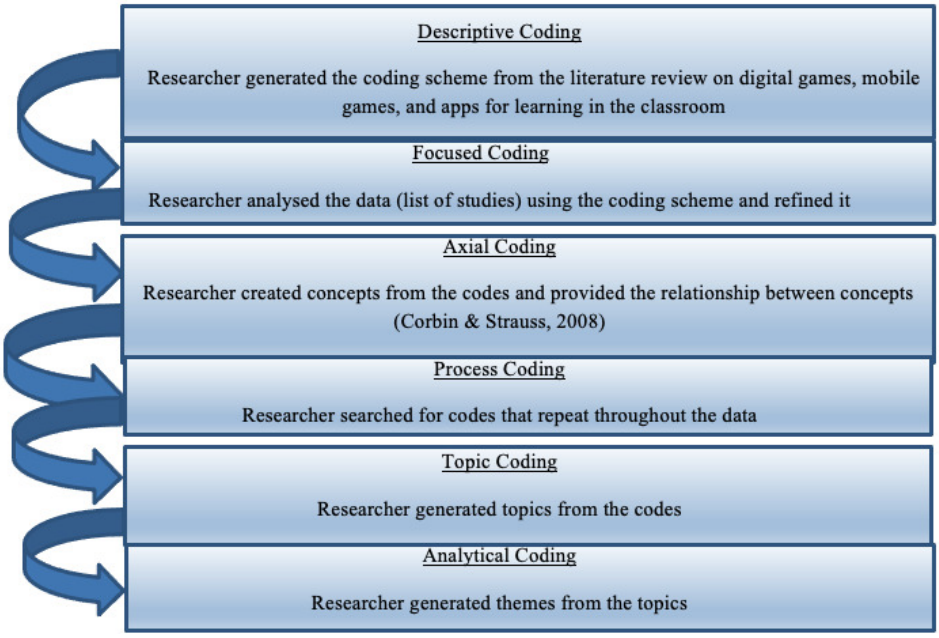

Figure 1. Qualitative coding process for the study

the classroom for teaching and learning of the students (Annamalai, 2019; Hmidani & Zareian, 2022; Poláková & Klímová, 2019). However, there has not been a recent review conducted on relevant studies that examined the apps' affordances for language learning. Second, the COVID-19 pandemic has accelerated the shift towards digital learning (Nilsberth et al., 2021; van der Spoel et al., 2020). Therefore, it is crucial to enable teachers to understand how apps can be used to promote language learning (Kaharuddin, 2020).

3.4. Data extraction and coding

After an agreement between the two researchers on the data extraction form, the following data were extracted from the selected articles in the systematic review and were coded. An example of the data extraction form used is shown in Table 1.

A second coder checked the coding and a consensus was reached in cases of disagreements in the coding. Figure 1 illustrates the coding process.

3.5. Synthesis of results

The heterogeneity of the selected studies in the systematic review did not enable a standard quantitative synthesis (i.e., metanalysis) to be conducted. Instead, the synthesis was narrative and a thematic analysis (Grant & Booth, 2009) was performed to extract large-scale patterns from the analysis of past studies. The summary tables in Appendix 1 (accessed at https://tinyurl.com/tewtapps) presented the synthesis results. During the synthesis process, the studies' results were critically discussed by the two researchers with the aim of identifying descriptive categories (i.e., types of apps for learning), to which the

Table 1. Example of the data extraction form used

Author/Year	Country/region	Research design	Sample size	Students' age	App's name	Desired learning outcome(s)	Learning gains	App's affordances
Çinar & Arı (2019)	Turkey	Pre- and Post-Test	71	14 to 16 years	*Quizlet*	Support vocabulary learning and increase positive attitudes	• Students' vocabulary scores increased significantly after four weeks of practice • Students retained vocabulary words after four weeks of practice • Positive attitudes	• Gamified elements • Instant feedback • "Drill-and-practice" learning
Montaner-Villalba (2019)	Spain	Pre- and Post-Test	24	14 to 15 years	*Quizlet*	Vocabulary acquisition in foreign language learning	• EFL vocabulary and writing improved significantly at post-test after using *Quizlet* for a year	• Gamified elements • Instant feedback • "Drill-and-practice" learning
Toy & Büyükkarci (2019)	Turkey	Pre- and Post-Test, survey	200	13 to 14 years	*Quizlet*	Examine effects of using Quizlet on EFL learners' vocabulary learning and perceptions	• Students' vocabulary learning for word recognition, comprehension, and reading skills increased after 8 weeks	• Gamified elements • Track progress • Customisable • "Drill-and-practice" learning

results (i.e., learning outcomes) could be assigned based on evident similarity (Munn et al., 2014). Uncertainties were resolved through a consensus-based decision.

3.6. Demographic and methodological summary of the included studies

Of the 38 articles, 13.2% were qualitative studies (n = 5), 26.3% were mixed methods (n = 10), while 60.5% reported quantitative analyses (n = 23). Most qualitative studies drew on interviews, analyses of students' artifacts, observations, and journals to present ethnographic or phenomenological case studies. The mixed methods studies used a combination of pre- and post-tests for assessing students' language learning, surveys, interviews, and observations. The quantitative studies included a mix of surveys and pre- and post-tests to assess students' language learning outcomes from using apps. Most of the studies (n = 6) were from Taiwan. The remaining studies were from Turkey (n = 5), Saudi Arabia (n = 3), Spain (n = 3), Colombia (n = 2), Indonesia (n = 2), Iran (n = 2), Malaysia (n = 2), United States (n = 2), Belgium (n = 1), Chile (n = 1), China (n = 1), Czech Republic (n = 1), Ecuador (n = 1), Germany (n = 1), Japan (n = 1), Korea (n = 1), Malta (n = 1), Palestine (n = 1), and the Netherlands (n = 1).

3.7. Quality of relevant studies

When the included article fulfilled all the inclusion criteria, we assessed it as a high-quality article. When the included article fulfilled between four to seven out of the eight inclusion criteria, we assessed it as a medium-quality article. When the included article fulfilled between one to three out of the eight inclusion criteria, we assessed it as a low-quality article. A total of 33 studies fulfilled all the inclusion criteria whereas five articles fulfilled seven out of the eight inclusion criteria (see Appendix 1). Two of the articles, including Çinar and Arı (2019) and Montaner-Villalba (2019), did not explicitly describe the apps' affordances that are relevant for the ESL/EFL learning outcomes and three articles, including Arman (2021), Cho (2021), and Kálecký (2016), were not peer-reviewed. Therefore, there are 33 high-quality articles and five medium-quality articles.

4. Findings

The thematic analysis was guided by the research question of the study, that is, to identify the typical learning gains reported in relation to specific categories of apps and discuss the affordances that typically characterised the types of apps.

4.1. Quiz apps support vocabulary acquisition

Quiz apps are student response systems that teachers use to create simple quizzes that students can take on their mobile devices. The quiz apps included Quizlet, Quizizz, Kahoot, Duolingo, Socrative, and Words&Birds. Twelve studies reported the use of quiz apps to promote language learning. In

particular, most studies reported that quiz apps can support vocabulary acquisition. A few studies also described learning gains in grammar learning (Rahayu & Purnawarman, 2019), reading development (Mulder et al., 2021), as well as improvement in students' attitude, confidence, and reception towards learning (Sercanoğlu et al., 2021).

Eight studies (Cho, 2021; Çinar & Arı, 2019; Dreyer, 2014; Guaqueta & Castro-Garces, 2018; Ludwig, 2018; Montaner-Villalba, 2019; Quiroz et al., 2021; Toy & Büyükkarci, 2019) reported that quiz apps provided positive gains in vocabulary learning for students. For example, Toy and Büyükkarci (2019) found that word recognition, comprehension, and reading for EFL learners' vocabulary learning improved after eight weeks of using the Quizlet app. Çinar and Arı (2019) also reported that EFL students' vocabulary scores and vocabulary retention test scores increased after four weeks of practice using the Quizlet app. Additionally, students' attitudes towards the English course became more positive after using the Quizlet app. Montaner-Villalba (2019) observed that EFL students who used Quizlet improved their vocabulary and writing in the post-test score after using Quizlet for an academic year. Cho (2021) also found that the Quizlet group outperformed the traditional group in vocabulary test scores after a treatment period of ten weeks and the students indicated that they had positive perceptions of Quizlet in studying target vocabulary.

Dreyer (2014) found that EFL students who used Quizlet over the course of 14 weeks scored higher in vocabulary and had less score variance than students in the non-Quizlet group. Additionally, frequent vocabulary review using Quizlet by the students several days before the test was found to improve their test scores. Guaqueta and Castro-Garces (2018) also reported that EFL students who used Duolingo and Kahoot over a six-month period had gained more vocabulary. Quiroz et al. (2021) found that EFL students who used Kahoot during two lessons a week for four weeks improved their English vocabulary knowledge. Ludwig (2018) also observed that EFL students' use of the Socrative app during a period of two weeks had a positive impact on students' overall vocabulary performance. However, it was noted that the multiple-choice-exercises used in the app may have contributed to students guessing the answers rather than remembering the correct translation of words.

Notwithstanding the positive reports, one study, Kálecký (2016), reported that when EFL students studied vocabulary using Quizlet, decontextualisation of word learning did not lead to vocabulary improvement. Nevertheless, he observed that Quizlet was helpful in eliminating spelling mistakes and could lead to better long-term retention when pictures were used for vocabulary learning in a contextualised manner.

Other than vocabulary learning, one study (i.e., Mulder et al., 2021) found quiz apps useful for students' reading development. There was only one study (i.e., Rahayu & Purnawarman, 2019) which investigated grammar learning through

quiz apps. Using a qualitative case study approach, Rahayu and Purnawarman (2019) found that EFL students' self-assessment using Quizizz improved their grammar understanding by helping them to identify their strengths and weaknesses when they reviewed the questions. However, they also observed that the majority of students were often distracted by the question timer when they intended to increase their rank in the gamified system by answering in a very short time.

The affordances of quiz apps which support vocabulary learning could be the gamified elements, such as a competitive leader board to track progress and provide automated and timely feedback on learning, as well as allowing for various customisation to personalise the digital play experience. Quiz apps also encourage a "drill-and-practice" learning strategy that could be effective for vocabulary acquisition. However, as noted by Kálecký (2016), such a decontextualised approach could be less effective, unless the words are introduced together with pictures and sentences in the game.

4.2. Puzzle apps support vocabulary and grammar learning

Puzzle apps are games that require problem-solving skills such as logic, pattern recognition, sequence solving, spatial recognition, and word completion. The puzzle apps included web-based mini games, interactive jigsaw puzzle, and card design game.

Eight studies reported that puzzle apps brought about learning gains in grammar and vocabulary learning. Mifsud et al. (2013) found both vocabulary and grammar improvements for ESL students who used the ClueFinders Reading Adventures app in their English lessons for a period of four weeks, with each daily session lasting 40 minutes. Shokri and Abdolmanafi-Rokni (2014) reported that EFL students' vocabulary, in terms of spelling performance, was enhanced after using the puzzle apps such as Hangman, Spelling bee, Fast hand, and Concentration games. Learning outcomes from using the apps indicated EFL students' improved recall and retention of spelling. Fariba et al. (2019) found that the digital game software SHAIEX, an adaptive hypermedia system for the teaching of languages at early ages, contributed to a significant effect on intermediate EFL learners' vocabulary learning.

Hao et al. (2019) found that middle school EFL students' overall vocabulary score increased after using the Detective ABC app for 40 minutes. They also found that EFL students' spelling, reading, and listening comprehension improved and they acquired opportunities for repeated oral practice. Calvo-Ferrer and Belda-Medina (2021) reported that ESL students who used new L2 words in the context of the online game Among Us in four sessions in two different weeks retained more vocabulary words than players who only encountered them. Cornillie et al. (2017) conducted a one-month study to examine students' L2 grammar practice using mini-games and reported that grammar practice using the mini-games helped students to develop

automaticity. Letchumanan and Tan (2012) involved their students in playing computer games for seven weeks to learn vocabulary. They found that there was a significant difference between the pre- and post-vocabulary test scores.

In addition to increased vocabulary and grammar learning, studies on puzzle apps have also reported learning gains in students' speaking skills. One study (Hwang et al., 2016) found that EFL students' verbal skills improved after using the interactive jigsaw puzzle and card design game on a mobile system during the learning activities in weeks one, two, and three. However, the statistical analysis indicated that although the average score for the use of games on the mobile system was significantly correlated with the verbal post-test scores, it had no significant correlation with the speaking improvement.

Similar to quiz apps, the affordances of puzzle apps include gamified elements such as points, badges, and leader boards, as well as the provision of instant feedback to learners, and the ability for teachers to track the learners' progress. In addition, the apps offer a competitive learning environment which motivates the students to practice with each repetitive play. Other affordances of the apps included user-friendliness, and ease of access, which are unsurprising as they are characteristics of all good apps (Tang & Hew, 2017).

4.3. Platform apps support reading and writing skills development

Platform apps are mobile or web-based learning management systems that are learner-centric and designed to promote self-directed learning. The platform apps used in the studies reviewed included Pixton, a platform designed to create comic strips online and learning management systems such as DynEd, Edmodo, Moodle, Virtual Room, WebQuest, Voki, and SeeSaw.

Ten studies reported that platform apps can bring about learning gains in reading and writing skills, which also included aspects of vocabulary and grammar. Tanduklangi et al. (2019) conducted four weeks of classes using Edmodo. The teacher created lesson scenarios, designed content and activities that were displayed through Edmodo and reported that students improved in reading and writing skills. Specific to reading skills, Chen et al. (2016) also reported improved oral reading fluency in EFL learners after using Moodle and the Qu-voice digital pen. Alpala and Peña (2014) reported an improvement in EFL students' writing through their text writing workshops using the Virtual Room app. Qaddumi (2021) conducted a study for 12 weeks where the experimental group was taught with Edmodo and reported that students showed learning gains in reading and writing. Beyond reading and writing, Yeşilbağ and Korkmaz (2021) found that students of the experimental group taught with Voki-supported activities for six weeks showed improvements in speaking. Shaikh et al. (2021) found that EFL students who used *DynEd* for eight weeks showed improvements in listening, writing, and reading skills.

Platform app use support both vocabulary and grammar learning. In Al-Kathiri's (2015) study, EFL students reported that Edmodo helped them to acquire new English vocabulary and the majority of students reported improved spelling and grammar. Al-Essa (2018) also reported a study where the experimental EFL group received grammar exercises through Edmodo for six weeks and showed improvement in their grammar competency. In Al-Harbi and Alshumaimeri's (2016) study, the experimental group learnt using videos on Edmodo for seven class periods (45 minutes) over six weeks. The mean English grammar score of the experimental group was higher than that of the control group despite not being statistically significant. In Cabrera et al.'s (2018) study, EFL students who received English lessons using Pixton for four months improved their EFL grammar and vocabulary.

Platform apps also develop critical thinking and digital literacies. Renau and Pesudo (2016) implemented WebQuest to promote ESL students' reading, speaking and writing skills, intercultural and digital competence, and critical thinking in the secondary ESL classroom. They reported that students developed increased critical thinking, cultural and digital competence, even as they improved in reading and writing skills.

Not all studies reported positive learning gains from the use of platform apps though. One study (i.e., K.-T. Wong et al., 2018) on the use of Moodle to supplement classroom instruction found no significant difference in ESL students' language learning gains even though positive effects were found on both learner autonomy and students' engagement as compared to the conventional learning. Arman's (2021) study on students' use of SeeSaw for interactive journalling also found that there was little improvement for students' writing self-efficacy, attitude toward writing, and writing performance.

The affordances of platform apps that support the development of reading and writing skills include hosting a range of relevant and relatable learning resources as well as rich multimedia content. Platform apps also offer interactive affordances to deepen students' engagement with the learning tasks and support collaborative learning with their peers online. Many platform apps also use analytics to provide feedback, through various forms of visualisations, to both teacher and students.

4.4. Augmented reality apps support increased engagement

Augmented reality apps layer computer-generated images on reality to create a composite view that augments the real world. The augmented reality apps in the studies reviewed are an augmented reality mystery app and a context-aware ubiquitous learning environment called the Handheld English Language Learning Organization (HELLO).

Studies on augmented reality apps reported that they increase students' engagement and interest towards language learning. Frazier (2017) designed an iPad app with integrated augmented reality technology for use in the ESL/EFL classroom over three days. The learning activity involved a mystery in an augmented reality crime room and students had to work in groups to solve the puzzle using the iPad app. Frazier (2017) reported that students developed critical thinking because they could analyse the meanings of the messages and correlate the mystery with the meaning of the objects through which the messages were conveyed.

Liu and Chu (2010) reported on the trial Handheld English Language Learning Organization (HELLO) to increase ESL/EFL learner's engagement. An eight-week experiment was conducted with one lesson per week. The students reported higher learning engagement and that the interaction with a virtual learning tutor developed their confidence in communication.

Augmented reality apps are relatively new, and the few studies on their use in the secondary English classroom tended to focus on the common outcome of learning engagement. This could be that the affordances of augmented reality apps engage the students' interest through the novelty of presenting the knowledge from the lesson materials in a virtual 3-dimension format. However, the learning gains in language development from the use of augmented reality apps remain to be affirmed with further studies.

4.5. Virtual reality apps support development of listening and speaking skills

Virtual reality apps offer an immersive experience through the wearing of headgear that simulate a virtual environment. Virtual reality apps used in the studies reviewed included virtual reality educational games, spherical video-based virtual reality, and a digital task-collaborative board game platform.

Given the emerging nature of virtual reality apps, only three studies on the use of these apps in the secondary school classroom were found. These studies reported that virtual reality apps support students' development of listening and speaking skills. Wu et al. (2014) reported that EFL students' communication skills were improved using digital board game with virtual reality affordances such as immersion and presence. Quantitative results indicated learning in terms of fluency and vocabulary acquisition. Qualitative results indicated that the virtual reality simulation provided context-relevant immersion and encouraged students' practice of listening and speaking skills as well as enabled them to foster intrinsic motivation. Chien et al. (2020) reported that EFL students developed stronger communication skills after using a spherical video-based virtual reality (SVVR) environment to situate students in authentic English-speaking contexts. Students in the virtual reality scenario teaching group made greater progress in vocabulary acquisition and speaking skills.

The affordances of virtual reality apps are that they offer an immersive virtual experience for the students and can simulate a context for communication practice. This is achieved with the use of virtual avatars, which reduce students' anxiety when they practise communication skills. As virtual reality apps are still nascent, the value of such apps for language learning is informed by just three studies, all attesting to its use to develop listening and speaking skills.

5. Discussion

We have summarised the key relations based on the findings from the studies reviewed. However, the learning gains are not limited by what is highlighted in the headlines of the earlier section. In terms of quantitative studies, it is important to note that only some studies reported significant language learning gains obtained from the differences between the pre-test, post-test, and control groups. In addition, the retention of learning, often not reported in the studies, is uncertain and may be low especially when the app is not used extensively or regularly. It is unrealistic to expect sustained learning gains from a one-time or infrequent usage of an app. It must be acknowledged that language learning cannot be meaningfully achieved within a short intervention period. For example, for students' reading comprehension and vocabulary to improve significantly, vocabulary learning had to be carried out in an incremental process where words must be met and used multiple times and reviewed over time to be truly learned (Schmitt, 1997).

In terms of the qualitative studies, it is also important to note that the reports on the learning gains were based on the perception of students and observations made by the researchers. The results could be triangulated across various data sources in some studies although the analysis and findings were usually context-bound. Regardless, these studies have been included in the systematic review, and their results reported in this paper provide a more holistic understanding of the range of studies conducted and the typical learning gains reported with specific categories of apps for learning.

It is also important, nonetheless, to acknowledge that it does not mean that the use of apps for learning will always bring about learning gains as there are other factors involved in the effective use of apps for learning. The first factor is the recognition that regardless of the quality of learning technologies and resources, it is ultimately decisive how they are used in support of the teacher's pedagogy that determines whether effective learning is achieved (Lim, 2021; Moseley et al., 1999; Oliver et al., 2007; Petko, 2012). In other words, it is the teacher, not the tools or materials, that makes the difference. As the OECD (2015) report on Students, Computers and Learning neatly concludes, "In the end, technology can amplify great teaching, but great technology cannot replace poor teaching" (p. 17). Notwithstanding, operating on the premise that we have competent teaching, it is useful to identify the "good" digital resources that can support the teacher's pedagogy and to bring about the desired language learning.

The second factor is the profile and readiness of students to engage in digital learning. For instance, students with more experience in using apps in out-of-school contexts do not need to learn how to use apps in the classroom and could therefore serve as mentors to other students to socialise them into the app literacy practices. The popular view of youths today as "digital natives" (Prensky, 2001) has been contested (Bennett et al., 2008) as access to technology and proficient digital skills amongst young people can differ across socio-economic status and cultures.

The third factor is the recognition that a classroom ecology is needed for effective integration of digital resources for teaching and learning (Hsu & Kuan, 2013; Martinez-Maldonado et al., 2015). This includes infrastructure such as a reliable internet connection, configurability of digital tools according to pedagogical intentions (Tchounikine, 2013), students' access to devices, as well as students' readiness habitualised through routine protocols on the use of digital tools for learning. Notwithstanding the limited discussion of these factors in this paper, they form the assumptions we have as we identify the value of specific categories of educational apps for certain types of learning that can be achieved in an environment of effective pedagogy and conducive classroom ecology.

While we have identified the learning associated with specific categories of educational apps, we acknowledge that this study has only reviewed a small sample size of 35 research studies that examined the use of apps for learning in the English classroom. As this review included only articles written in English, important studies that have been published in other languages may have been excluded. By only using online search engines for the systematic search of relevant papers, we could have missed important findings published in hardcopy books and edited volumes.

Most of our studies in the systematic review, except for Cabrera et al. (2018), Cho (2021), Cornillie et al. (2017), Mulder et al. (2021), Mifsud et al. (2013), Shaikh et al. (2021), Toy & Büyükkarci (2019), and Wong et al. (2018), included research designs with a small sample size. While they provide us with detailed descriptions of teachers using apps for learning in the classroom across a variety of contexts in different countries/regions, their findings on students' language learning outcomes are not generalisable due to a small sample size. Future research on apps for learning in the English classrooms should be conducted with a larger sample size to improve the generalisability of the findings.

6. Conclusion

Through our systematic review, we have identified the typical learning gains reported from the use of various types of educational apps. The systematic review highlights the studies on this topic that have been done in the last decade. It calls for further studies to build on existing work and also surfaces gaps in our current understandings of the value of specific apps for different

types of learning and shows the need for more studies on emerging technologies, such as the use of augmented reality and virtual reality apps to examine their affordances in relation to the learning outcomes for students. It is hoped that our research can help teachers make sense of the studies on apps for learning and to guide their selection and appropriate use of educational apps for language learning in the L2 English classroom.

Acknowledgement

This study was funded by the SUG-NAP (Start-up Grant for New Assistant Professors) administered by National Institute of Education, Nanyang Technological University, Singapore. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the university.

Authorship

Lim, F.V.: Conceptualization; Funding acquisition; Project administration; Resources; Supervision; Validation; Visualization; Writing - review & editing.

Toh, W.: Data curation; Formal analysis; Investigation; Methodology; Validation; Roles/Writing - original draft. All authors approve final version of the article.

Declaration of interest

Authors declare no competing interest.

Submitted: February 20, 2024 EET

REFERENCES

Al-Essa, N. S. (2018). *The impact of using Edmodo as a blended learning medium on promoting Saudi EFL female secondary school students' English grammar* [Master's thesis, Al Imam Muhammad ibn Saud Islamic University]. https://doi.org/10.24093/awej/th.221

Al-Harbi, S. S., & Alshumaimeri, Y. A. (2016). The flipped classroom impact in grammar class on EFL Saudi secondary school students' performances and attitudes. *English Language Teaching*, *9*(10), 60–80. https://doi.org/10.5539/elt.v9n10p60

Al-Kathiri, F. (2015). Beyond the classroom walls: Edmodo in Saudi secondary school EFL instruction, attitudes and challenges. *English Language Teaching*, *8*(1), 189–204. https://doi.org/10.5539/elt.v8n1p189

Alpala, D. P. O., & Peña, N. M. (2014). A virtual room to enhance writing skills in the EFL class. *HOW Journal*, *21*(1), 62–81. https://doi.org/10.19183/how.21.1.15

Annamalai, N. (2019). Using Whatsapp to extend learning in a blended classroom environment. *Teaching English with Technology*, *19*(1), 3–20. https://files.eric.ed.gov/fulltext/EJ1204549.pdf

Arman, N. M. (2021). *The impact of interactive journaling on seventh-grade students' writing self-efficacy, writing performance, and attitudes towards writing* [Unpublished PhD dissertation, University of South Carolina]. https://scholarcommons.sc.edu/cgi/viewcontent.cgi?article=7161&context=etd

Bennett, S., Maton, K., & Kervin, L. (2008). The 'digital natives' debate: A critical review of the evidence. *British Journal of Educational Technology*, *39*(5), 775–786. https://doi.org/10.1016/10.1111/j.1467-8535.2007.00793.x

Brand, J., & Kinash, S. (2010). Pad-agogy: A quasi-experimental and ethnographic pilot test of the iPad in a blended mobile learning environment. In C. H. Steel, M. J. Keppell, P. Gerbic, & S. Housego (Eds.), *Curriculum, technology & transformation for an unknown future. Proceedings of the Ascilite Conference* (pp. 147–151). Ascilite. https://www.researchgate.net/publication/47929709_Pad-agogy_A_quasi-experimental_and_ethnographic_pilot_test_of_the_iPad_in_a_blended_mobile_learning_environment

Cabrera, P., Castillo, L., González, P., Quiñónez, A., & Ochoa, C. (2018). The impact of using Pixton for teaching grammar and vocabulary in the EFL Ecuadorian context. *Teaching English with Technology*, *18*(1), 53–76.

Calvo-Ferrer, J. R., & Belda-Medina, J. (2021). The effect of multiplayer video games on incidental and intentional L2 vocabulary learning: The case of Among Us. *Multimodal Technologies and Interaction*, *5*(12), 1–16. https://doi.org/10.3390/mti5120080

Carrier, M., Damerow, R. M., & Bailey, K. M. (2017). *Digital Language Learning and Teaching - Research, Theory, and Practice*. Routledge.

Chen, C.-M., Tan, C.-C., & Lo, B.-J. (2016). Facilitating English-language learners' oral reading fluency with digital pen technology. *Interactive Learning Environments*, *24*(1), 96–118. https://doi.org/10.1080/10494820.2013.817442

Cheung, C. K., & Slavin, R. E. (2012). How features of educational technology applications affect student reading outcomes: A meta-analysis. *Educational Research Review*, *7*(3), 198–215. https://doi.org/10.1016/j.edurev.2012.05.002

Chien, S.-Y., Hwang, G.-J., & Jong, M. S.-Y. (2020). Effects of peer assessment within the context of spherical video-based virtual reality on EFL students' English-speaking performance and learning perceptions. *Computers & Education*, *146*, 1–20. https://doi.org/10.1016/j.compedu.2019.103751

Cho, H. (2021). *Quizlet in the EFL classroom: Enhancing vocabulary acquisition of Korean middle school students* [Unpublished PhD dissertation]. Alliant International University.

Çinar, İ., & Arı, A. (2019). The effects of Quizlet on secondary school students' vocabulary learning and attitudes towards English. *Asian Journal of Instruction*, *7*(2), 60–73. https://www.proquest.com/docview/2468395993/fulltextPDF/23B7D21FE23F4368PQ/1?accountid=28158

Cooper, H. (2016). *Research Synthesis and Meta-Analysis: A Step-by-Step Approach* (5th ed.). SAGE Publications.

Cornillie, F., Van den Noortgate, W., Van den Branden, K., & Desmet, P. (2017). Examining focused L2 practice: From in vitro to in vitro. *Language Learning & Technology*, *21*(1), 121–145.

Dreyer, J. (2014). The effect of computer-based self-access learning on weekly vocabulary test scores. *Studies in Self-Access Learning Journal*, *5*(3), 217–234. https://doi.org/10.37237/050303

Fariba, R. E., Mohammad Reza, R. T., & Mehrnnoosh, H. (2019). The impact of digital games on intermediate EFL learners' vocabulary improvement. *International Journal of Research Studies in Language Learning*, *8*(1), 29–38. https://doi.org/10.5861/ijrsll.2019.3013

Frazier, E. E. (2017). Enhancing learner experience through augmented reality in high school. *PanSIGJournal*, *33*(2017), 33–39.

Gangaiamaran, R., & Pasupathi, M. (2017). Review on use of mobile apps for language learning. *International Journal of Applied Engineering Research*, *12*(21), 11242–11251. https://www.ripublication.com/ijaer17/ijaerv12n21_102.pdf

Grant, M. J., & Booth, A. (2009). A typology of reviews: An analysis of 14 review types and associated methodologies. *Health Information & Libraries Journal*, *26*(2), 91–108. https://doi.org/10.1111/j.1471-1842.2009.00848.x

Guaqueta, C. A., & Castro-Garces, A. Y. (2018). The use of language learning apps as a didactic tool for EFL vocabulary building. *English Language Teaching*, *11*(2), 61–71. https://doi.org/10.5539/elt.v11n2p61

Gustafsson, U. (2021). Taking a step back for a leap forward: Policy formation for the digitalisation of schools from the views of Swedish national policymakers. *Education Inquiry*, *12*(4), 329–346. https://doi.org/10.1080/20004508.2021.1917487

Hao, Y., Lee, K. S., Chen, S.-T., & Sim, S. C. (2019). An evaluative study of a mobile application for middle school students struggling with English vocabulary learning. *Computers in Human Behavior*, *95*, 208–216. https://doi.org/10.1016/j.chb.2018.10.013

Hmidani, T., & Zareian, N. (2022). Mobile-mediated interactional feedback (MMIF) effect on Iranian learners' acquisition of English articles. *Teaching English with Technology*, *22*(1), 40–61. https://files.eric.ed.gov/fulltext/EJ1327330.pdf

Hsu, S., & Kuan, P.-Y. (2013). The impact of multilevel factors on technology integration: The case of Taiwanese grade 1–9 teachers and schools. *Educational Technology Research and Development*, *61*(1), 25–50. https://doi.org/10.1007/s11423-012-9269-y

Hwang, W.-Y., Shih, T. K., Ma, Z.-H., Shadiev, R., & Chen, S.-Y. (2016). Evaluating listening and speaking skills in a mobile game-based learning environment with situational contexts. *Computer Assisted Language Learning*, *29*(4), 639–657. https://doi.org/10.1080/09588221.2015.1016438

Kaharuddin, A. (2020). Contributions of technology, culture, and attitude to English learning motivation during COVID-19 outbreaks. *Systematic Reviews in Pharmacy*, *11*(11), 76–84.

Kálecký, R. (2016). *Quizlet vs. Vocabulary Notebook: The Impact of Different Methods of Storing and Revising Vocabulary on Students' Progress, Retention and Autonomy* [Master's diploma thesis]. Masaryk University.

Letchumanan, K., & Tan, B. H. (2012). Using computer games to improve secondary school students' vocabulary acquisition in English. *Pertanika Journal of Social Science and Humanities*, *20*(4), 1005–1018.

Lim, F. V. (2021). *Designing Learning with Embodied Teaching: Perspectives from Multimodality*. Routledge.

Lim, F. V., & Toh, W. (2022). Considerations on the curation of educational apps for digital play and learning. *Contemporary Educational Technology*, *14*(3), ep366. https://doi.org/10.30935/cedtech/11809

Liu, T.-Y., & Chu, Y.-L. (2010). Using ubiquitous games in an English listening and speaking course: Impact on learning outcomes and motivation. *Computers & Education*, *55*(2), 630–643. https://doi.org/10.1016/j.compedu.2010.02.023

Ludwig, C. (2018). Using vocabulary apps to enhance students' vocabulary knowledge. *SiSAL Journal*, *9*(2), 306–323. https://doi.org/10.37237/090305

Macaro, E., Handley, Z., & Walter, C. (2012). A systematic review of CALL in English as a second language: Focus on primary and secondary education. *Language Teaching*, *45*(1), 1–43. https://doi.org/10.1017/s0261444811000395

Martinez-Maldonado, R., Clayphan, A., & Kay, J. (2015). Deploying and visualising teacher's scripts of small group activities in a multi-surface classroom ecology: a study in-the-wild. *Computer Supported Cooperative Work (CSCW)*, *24*(2–3), 177–221. https://doi.org/10.1007/s10606-015-9217-6

Mifsud, C. L., Vella, R., & Camilleri, L. (2013). Attitudes towards and effects of the use of video games in classroom learning with specific reference to literacy attainment. *Research in Education*, *90*(1), 32–52. https://doi.org/10.7227/rie.90.1.3

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *BMJ*, *339*, b2535. https://doi.org/10.1136/bmj.b2535

Montaner-Villalba, S. (2019). The use of Quizlet to enhance vocabulary in the English language classroom. In F. Meunier, J. Van de Vyver, L. Bradley, & S. Thouësny (Eds.), *CALL and complexity – short papers from EUROCALL 2019* (pp. 304–309). Research-publishing.net. https://doi.org/10.14705/rpnet.2019.38.1027

Moseley, D., Higgins, S., Bramald, R., Hardman, F., Miller, J., Mroz, M., Tse, H., Newton, D., Thompson, I., Williamson, J., Halligan, J., Bramald, S., Newton, L., Tymms, P., Henderson, B., & Stout, J. (1999). *Ways forward with ICT: Effective pedagogy using information and communications technology for literacy and numeracy in primary schools* (pp. 1–129). University of Newcastle. https://files.eric.ed.gov/fulltext/ED458652.pdf

Mulder, E., van de Ven, M., Segers, E., Krepel, A., de Bree, E. H., van der Maas, H., de Jong, P. F., & Verhoeven, L. (2021). Serious game-based word-to-text integration intervention effects in English as a second language. *Contemporary Educational Psychology*, *65*, 1–13. https://doi.org/10.1016/j.cedpsych.2021.101972

Munn, Z., Tufanaru, C., & Aromataris, E. (2014). JBI's systematic reviews: Data extraction and synthesis. *The American Journal of Nursing*, *114*(7), 49–54. https://doi.org/10.1097/01.naj.0000451683.66447.89

Nilsberth, M., Liljekvist, Y., Olin-Scheller, C., Samuelsson, J., & Hallquist, C. (2021). Digital teaching as the new normal? Swedish upper secondary teachers' experiences of emergency remote teaching during the COVID-19 crisis. *European Educational Research Journal*, *20*(4), 442–462. https://doi.org/10.1177/14749041211022480

OECD. (2015). *Students, Computers and Learning: Making the Connection*. OECD Publishing. https://doi.org/10.1787/9789264239555-en

Oliver, R., Harper, B., Wills, S., Agostinho, S., & Hedberg, J. (2007). Describing ICT-based learning designs that promote quality learning outcomes. In H. Beetham & R. Sharpe (Eds.), *Rethinking Pedagogy for a Digital Age: Designing and delivering e-learning* (pp. 64–80). Routledge.

Olofsson, A. D., & Lindberg, J. O. (2021). A glimpse of a Nordic model? Policy and practice in the digitalisation of the K-12 school and teacher education in Denmark, Finland, Norway and Sweden: Editorial introduction. *Education Inquiry*, *12*(4), 311–316. https://doi.org/10.1080/20004508.2021.1988451

Paez, A. (2017). Gray literature: An important resource in systematic reviews. *Journal of Evidence-Based Medicine*, *10*(3), 233–240. https://doi.org/10.1111/jebm.12266

Perrotta, C., Featherstone, G., Aston, H., & Houghton, E. (2013). *Game-based Learning: Latest Evidence and Future Directions* (NFER Research Programme: Innovation in Education). NFER.

Petko, D. (2012). Teachers' pedagogical beliefs and their use of digital media in classrooms: Sharpening the focus of the 'will, skill, tool' model and integrating teachers' constructivist orientations. *Computers & Education*, *58*(4), 1351–1359. https://doi.org/10.1016/j.compedu.2011.12.013

Poláková, P., & Klímová, B. (2019). Mobile technology and Generation Z in the English language classroom—A preliminary study. *Education Sciences*, *9*(3), 1–11. https://doi.org/10.3390/educsci9030203

Prensky, M. (2001). Digital Natives, Digital Immigrants Part 1. *On the Horizon*, *9*(5), 1–6. https://doi.org/10.1108/10748120110424816

Qaddumi, H. A. (2021). A study on the impact of using Edmodo on students' achievement in English language skills and retention. *Education and Information Technologies*, *26*(5), 5591–5611. https://doi.org/10.1007/s10639-021-10510-6

Quiroz, M. F., Gutiérrez, R., Rocha, F., Valenzuela, M. P., & Vilches, C. (2021). Improving English vocabulary learning through Kahoot!: A quasi-experimental high school experience. *Teaching English with Technology*, *21*(2), 3–13.

Rahayu, I. S. D., & Purnawarman, P. (2019). The use of Quizizz in improving students' grammar understanding through self-assessment. In *Eleventh Conference on Applied Linguistics (CONAPLIN 2018)* (pp. 102–106). Atlantis Press. https://doi.org/10.2991/conaplin-18.2019.235

Renau, M. L., & Pesudo, M. (2016). Analysis of the implementation of a WebQuest for learning English in a secondary school in Spain. *International Journal of Education and Development Using Information and Communication Technology*, *12*(2), 26–49. https://files.eric.ed.gov/fulltext/EJ1111476.pdf

Schaefer, H. R., & Myers, J. L. (2017). Guidelines for performing systematic reviews in the development of toxicity factors. *Regulatory Toxicology and Pharmacology*, *91*, 124–141. https://doi.org/10.1016/j.yrtph.2017.10.008

Schmitt, N. (1997). *Vocabulary Learning Strategies: Description, Acquisition and Pedagogy*. Cambridge University Press.

Sercanoğlu, M., Bolat, Y. İ., & Göksu, İ. (2021). Kahoot! as a gamification tool in vocational education: More positive attitude, motivation and less anxiety in EFL. *Journal of Computer and Education Research*, *9*(18), 682–701. https://doi.org/10.18009/jcer.924882

Shaikh, G., Koçak, Ö., & Göksu, İ. (2021). Does DynEd affect students' attitudes and language skills in EFL? A case study. *Teaching English with Technology*, *21*(1), 75–93.

Shokri, H., & Abdolmanafi-Rokni, S. J. (2014). The effect of using educational computer games on recall and retention of spelling in Iranian EFL learners. *International Journal of Applied Linguistics & English Literature*, *3*(6), 169–175. https://doi.org/10.7575/aiac.ijalel.v.3n.6p.169

Tanduklangi, A., Lio, A., & Alberth, A. (2019). Classroom action research in teaching English for senior high school students through blended learning in Kendari of Indonesia. *Journal of E-Learning and Knowledge Society*, *15*(1), 169–182. https://doi.org/10.20368/1971-8829/1579

Tang, Y., & Hew, K. F. (2017). Is mobile instant messaging (MIM) useful in education? Examining its technological, pedagogical, and social affordances. *Educational Research Review*, *21*, 85–104. https://doi.org/10.1016/j.edurev.2017.05.001

Tchounikine, P. (2013). Clarifying design for orchestration: orchestration and orchestrable technology, scripting and conducting. *Computers & Education*, *69*(1), 500–503. https://doi.org/10.1016/j.compedu.2013.04.006

Toy, F., & Büyükkarci, K. (2019). The effects of Quizlet on foreign language learners' vocabulary learning success and perceptions. *I-Manager's Journal of Educational Technology*, *16*(3), 44–60. https://doi.org/10.26634/jet.16.3.16450

Vaish, V., & Towndrow, P. A. (2010). Multimodal literacy in language classrooms. In N. H. Hornberger & S. L. McKay (Eds.), *Sociolinguistics and Language Education* (pp. 317–346). Multilingual Matters. https://doi.org/10.21832/9781847692849-014

van der Spoel, I., Noroozi, O., Schuurink, E., & van Ginkel, S. (2020). Teachers' online teaching expectations and experiences during the Covid19-pandemic in the Netherlands. *European Journal of Teacher Education*, *43*(4), 623–638. https://doi.org/10.1080/02619768.2020.1821185

Wong, K.-T., Hwang, G.-J., Goh, S. C., & Arrif, S. K. (2018). Effects of blended learning pedagogical practices on students' motivation and autonomy for the teaching of short stories in upper secondary English. *Interactive Learning Environments*, *28*(4), 512–525. https://doi.org/10.1080/10494820.2018.1542318

Wong, L.-H., & Looi, C.-K. (2010). Vocabulary learning by mobile-assisted authentic content creation and social meaning-making: two case studies. *Journal of Computer Assisted Learning*, *26*(5), 421–433. https://doi.org/10.1111/j.1365-2729.2010.00357.x

Wu, C.-J., Chen, G.-D., & Huang, C.-W. (2014). Using digital board games for genuine communication in EFL classrooms. *Educational Technology Research and Development*, *62*(2), 209–226. https://doi.org/10.1007/s11423-013-9329-y

Yeşilbağ, S., & Korkmaz, Ö. (2021). The effect of Voki application on students' academic achievements and attitudes towards English course. *Education and Information Technologies*, *26*(1), 465–487. https://doi.org/10.1007/s10639-020-10264-7