Joanna Pitura
University of the National Education
Commission, Krakow
https://orcid.org/0000-0003-1659-1995
joanna.pitura@up.krakow.pl

Anna Turula
University of the National
Education Commission, Krakow
https://orcid.org/0000-0003-3086-8591
anna.turula@up.krakow.pl

Sabina Nowak
University of the National Education
Commission, Krakow
https://orcid.org/0000-0001-9761-8365
sabina.nowak@up.krakow.pl

Joanna Jakubik
University of the National
Education Commission, Krakow
https://orcid.org/0000-0003-1482-4353
joanna.jakubik2@up.krakow.pl

Yuliya Asotska-Wierzba
University of the National
Education Commission, Krakow
https://orcid.org/0000-0002-6045-7091
yuliya.asotska-wierzba@up.krakow.pl

Mapping the landscape of serious digital language learning games: Towards a Serious-Digital Game-based Language Learning (S-DGBLL) teacher competence framework

Serious games have been gaining increasing attention in the field of digital game-based language learning (DGBLL). However, to incorporate serious games into foreign/second language (L2) classrooms, teachers must be prepared with requisite knowledge and skills. To this end, following PRISMA guidelines, this study conducted a systematic review of research on serious digital language learning games with two aims. The main objective of this study was to map out the issues pertaining to serious digital language learning games that are

currently being addressed in empirical studies in the DGBLL area. The secondary objective was to leverage the findings to inform the development of a research-based framework of expertise for L2 teachers intending to incorporate serious digital game-based language learning (S-DGBLL) in their classrooms. Fifty articles published until April 2023 were analysed in terms of research methodologies, pedagogical characteristics, game design, and outcomes. The findings have implications for the proposal of an S-DGBLL teacher competence framework.

Keywords: serious games, digital game-based language learning, systematic literature review

Słowa kluczowe: poważne gry, cyfrowa nauka języka w oparciu o gry, systematyczny przegląd literatury

1. Introduction

Peterson and Jabbari (2023) emphasize two needs they describe as pressing in the field of computer assisted language learning: rigorous evaluative work and expanding areas such as digital game-based language learning (DGBLL). This is particularly important because, as noted by Blume (2020), the popularity of DGBLL in foreign/second (L2) language education should and could be more widespread. However, a timely question is what L2 teachers should know and understand about DGBLL to improve their teaching. Accordingly, the present paper is an attempt to meet the needs delineated above, with special regard to contributing to the development of language teachers' competences in DGBLL by focusing on serious digital language learning games through the "review work" postulated.

With the rapidly growing body of literature describing the development and application of serious games in the process of both formal and informal L2 learning, the need for regular and rigorous reviewing of research is unquestioned. This need is satisfied by the number of researchers–systematic reviews are becoming a popular scientific genre. The evaluative work their authors embark upon varies in its rationale, aims and specificity. Some aspects of the DGBLL research, like methodologies and outcomes, attract more attention while others, the broadly understood pedagogy of game-based language education, remain more neglected. Besides this, there are differences and a certain lack of balance in the rationale with which systematic reviews are undertaken. All systematic reviews aim to define the status quo of DGBLL as such, or a chosen aspect of it. A much smaller number of these evaluative works translate the findings into models, strategies and recommendations. What is notably lacking in this landscape are well-defined teacher competences necessary for effective game-based language

teaching. This gap presents a barrier to the adoption of serious games in L2 education.

Accordingly, following PRISMA guidelines (Page et al., 2021), this study sets out to conduct a systematic review of research on serious digital language learning games with two aims. The main objective is to map out the issues pertaining to serious digital game-based language learning (S-DGBLL) that are currently addressed in empirical studies in the DGBLL area. The secondary objective is to leverage the findings to inform the development of a research-based preliminary framework of S-DGDLL expertise for L2 teachers.

2. Literature review

2.1. Serious Digital Game-Based Language Learning –
theoretical considerations

Game-Based Learning (GBL) is "the union of educational learning theories, course curricula, and digital gameplay with the goal of enhancing the learning experience" and the concept of GBL is "fun learning through doing/playing and specifically designed, structured game learning materials which can stimulate the development of thinking skills and self-learning" (Dahalan et al., 2023: 6). To provide theoretical justification for the application of this approach in the area of L2 learning and teaching, Reinhardt (2019) outlines four key perspectives. First, a structural-behaviorist viewpoint acknowledges that L2 learning can occur through translating and understanding game language, being exposed to language, receiving positive and negative reinforcement in the form of rewards and penalties, etc. A psycho-cognitive approach recognises that L2 learning can take place via opportunities for language production, immersion in game content made possible by in-game narratives, activities, and discourses, as well as by observing language generated by other players, among others. A socially informed viewpoint posits that L2 learning can happen through learners' participation in social practices such as learning how to play, acting out roles as a player or character in a game, interacting with game discourses, and negotiating meaning with other players. Lastly, an ecological viewpoint holds that L2 language usage and learning are complex, systemic, non-linear, and emergent. The idea of affordance is offered within this perspective, denoting "a potential for action made available when an actor and ecological design align" in a game (ibid.).

The concept of "serious games" has been defined in various ways. Zyda (2005: 26) defines them as "a mental contest, played with a computer in accordance with specific rules, that uses entertainment to further government

or corporate training, education, health, public policy, and strategic communication objectives." While Michael and Chen (2005) acknowledge the entertainment aspect, they emphasise that education is their primary goal. Serious games can also serve as a means of advertising, political campaigning, or social justice awareness raising (Reinhardt, 2019). They have been applied in various sectors, taking the form of educational games, simulation games, news games, roleplaying games (RPGs), pervasive games, etc. (Maugard, 2019).

Serious games need to be distinguished from other instructional approaches that incorporate games in the field of L2 learning and teaching. Reinhardt (2019: 8–11) offers a helpful framework for understanding the difference. Specifically, his framework encompasses three categories: game-based, game-enhanced, and game-informed L2 learning and teaching. Game-based L2 learning and teaching utilizes educational games specifically designed for language learning purposes. Commercial, off-the-shelf games designed for entertainment rather than L2 education are used in game-enhanced L2 learning and instruction. Lastly, game-informed L2 learning and teaching involves incorporating game elements such as levels, leaderboards, and badges to gamify tasks, lessons, or curricula. Therefore, game-based L2 learning and teaching is a manifestation of Game-Based Learning (GBL), with the focus on serious games designed for L2 learning.

The idea of using games in education is not new. What has changed over the years is the adaptation of the concept and practice in various areas–such as language education, hence GBLL (Game-Based Language Learning), as well as the use of ICT to enhance the experience, resulting in DG-BLL (Digital Game-Based Language Learning). Both sub-areas of GBL have enjoyed considerable interest expressed in numerous publications (Peterson, Jabbari, 2023). As for the research foci (Peterson et al., 2020), studies to-date predominantly focus on the vocabulary learning potential of such games (cf. Chen, Hsu, 2020; Lai, Chen, 2023) as well as affective factors related to them (see also Reinders, Wattana, 2015). Other–and recommendable–foci or research highlights (cf. Peterson, Jabbari, 2023: 5–9) include: the role of theory, or studies drawing on developments in SLA research; advancing research through developments in the methodology of the studies undertaken in the area; evaluation of empirical studies through review and analysis; replication studies; investigating the role of the educator; ways games are integrated into curricula, as well as game creation and gaming in informal contexts. Embedded in this context is the focus of this review paper: serious digital game-based language learning (S-DGBLL).

2.2. Systematic reviews in the field of DGBLL

DGBLL has been subject to several systematic reviews, carried out regularly since the onset of the practice and the use of (and research into) serious digital games in the language classroom. This paper looks into eight such papers (Acquah, Katz, 2020; Almaki et al., 2023; Borona et al., 2018; Dhimolea et al., 2022; Hung et al., 2018; Naul, Liu, 2020; Weng, Chiu, 2023; Zhai, Wibowo, 2023).

The systematic reviews vary in how they can be situated in the DGBLL context. Two of the reviews–Hung et al. (2018) and Acquah and Katz (2020)– are what can be called *hard-DGBLL* reviews, concentrating on articles which define digital game-based language learning as their context. Several other reviews can be labelled *soft-DGBLL.* Some of them have a broader focus, on DGBL, with the inclusion of DGBLL, but not a main emphasis on it. This is the case of the review by Almaki et al. (2023), who reviewed 24 research outcomes published in the years 2016–2021 pertaining to the effectiveness of game-based learning in primary and secondary schools in all subjects, including languages, but not specifically focussing on this area. Other *soft* reviews include ones in which there is an insight into firstly, a selected aspect of DGBLL, or where DGBLL itself is offered a marginal focus. Naul and Liu's (2020) paper is an example of the former. It concentrates on narrative in serious games, in the evaluation of 44 papers to-date. Good examples of the latter are the review by Borona et al. (2018), where the main focus is 3D MultiUser Virtual Environments with educational language learning games, such as RPG, as one of the many aspects of the 3DMUVE papers they analyse; and the review by Dhimolea et al. (2022), in which the positive effect of Virtual Reality (VR) games on learners' pronunciation is just one of the many threads in a review of High-Immersion VR for Language Learning. The paper by Weng and Chiu (2023) is a very similar case. Finally, the DGBLL focus may also be indirect, as in the review by Zhai and Wibowo (2023) which concentrates solely on AI dialogue systems.

The aims of these studies are mostly cognitive in nature, i.e. the need to define the status quo. As a result, the authors (Acquah, Katz, 2020; Almaki et al., 2023; Borona et al., 2018; Dhimolea et al., 2022; Hung et al., 2018; Naul, Liu, 2020; Zhai, Wibowo, 2023) review studies in defined research periods from the perspective of discovering how game-based learning affects education, in terms of learning outcomes, as well as affective, social and behavioural factors which influence learning, to develop a better understanding of GBL as a method. Some studies go beyond "a need to understand," towards suggesting practical applications of the knowledge gained from the review, such as proposing a conceptual model for understanding how learners' motivation and computational-thinking skills can benefit from games

with a view to improving game implementation (Almaki et al., 2023); identifying a comprehensive set of evaluation dimensions (Zhai, Wibowo, 2023); or recommending game design strategies (Naul, Liu, 2020).

Typical foci of the reviews include the research methodologies of the evaluated studies, the languages taught, and the broadly understood learning outcomes (the majority of the studies). Other selected foci are: game features (Naul, Liu, 2020); game contexts, including learning environment and pedagogy (Acquah, Katz, 2020; Weng, Chiu, 2023; Zhai, Wibowo, 2023); limitations of the studies, and recommendations they make (Almaki et al., 2023; Zhai, Wibowo, 2023); and learner characteristics (Borona et al., 2018). What is important for the present paper, is that all the reviews investigate game genres /types. However, divisions into game-based and game-enhanced, with separation of serious games from commercial off-the-shelf (COTS) games are not systematically present. For example, one of the two *hard-DGBLL* reviews mentions studies dealing with "commercial off-the-shelf (COTS), freely available, and researcher-developed or custom-built games, on the basis of their availability to users [and also] … mixed or varied games used to code any studies that adopted more than one type of digital game" (Hung et al., 2018: 93).

Finally, the reviews are rather uniform as regards their methodologies and procedures. Most studies rely on the PRISMA statement (Acquah, Katz, 2020; Almaki et al., 2023; Dhimolea et al., 2022; Weng, Chiu, 2023; Zhai, Wibowo, 2023). Others (Borona et al., 2018; Hung et al., 2018; Naul, Liu, 2020) cite previously published work, that they treat as the source for standards of research. All the studies follow stages of identification of the sample, screening; eligibility decisions; selection and the process is carried out based on pre-established protocols (including keywords and inclusion/exclusion criteria). The research is performed by multiple researchers, and may include inter-rater validation.

The systematic reviews performed to-date give informative insights into the status quo of game-based learning research, with particular focus on research methodologies and learning outcomes. However, as their aim was to stake out territory, they seem to neglect points that are of interest to teacher training for DGBLL. The most interesting of the foci to date seem to be closer insights into the broadly understood pedagogies and game design strategies presented in the DGBLL studies. There are studies that investigate these aspects, but they usually concentrate on very particular aspects of GBL, such as AI, or narratives in games. The present study aims to fill these gaps by concentrating solely on studies of the design or use of serious digital language learning games. More specifically, the main objective is to map out issues pertaining to S-DGBLL that are currently addressed in empirical studies in the DGBLL area. The secondary objective is to leverage the

findings to inform the development of a preliminary, research-based framework of S-DGBLL expertise for L2 teachers. The following research questions guided the present investigation:

1. What methodologies have been employed to study serious digital language learning games?
2. What are the pedagogical characteristics of serious digital language learning games?
3. What are the design characteristics of serious digital language learning games?
4. What are the key outcomes of studies of serious digital language learning games?

3. Method

The present study adopted systematic review methodology and was guided by the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement (Page et al., 2021). The research activities included the following phases: (1) defining eligibility criteria, (2) obtaining the relevant papers, (3) screening and selecting the papers in accordance with the eligibility criteria, and (4) extracting and analysing data in order to answer the research questions.

To be included in the review, papers needed to be published in English as full-text articles, or, if they had been cited at least once, dissertations. In addition, it was necessary for papers to focus on foreign and/or second language learning and teaching; to report on the design, or use of a digital serious language game; and to provide empirical data. All papers published up until the date of the review were included.

Papers were not considered eligible when they were not written in English, when they concentrated on first or heritage languages, and when they were published as literature reviews, commentaries, theoretical studies, conceptual papers, editorials, opinion articles, uncited dissertations, or conference abstracts. Papers were excluded when they did not deal with serious games for foreign or second language learning and teaching, i.e. when they reported on gamification research, analogue game design and/or implementation, commercial games not designed for language learning purposes, or serious games not designed for language learning purposes.

The research team conducted the search in April 2023. To obtain the relevant papers, Google Scholar, Science Direct, Springer Link, Taylor & Francis, Scopus, Directory of Open Access Journals, ACM Digital Library, IEE-Explore, and Web of Science were browsed. Papers were also searched

for by examining journal websites and reference lists of relevant publications, and by using the "Cited by" function in Google Scholar (cf. Dhimolea et al., 2022). We carried out the search by using three sets of keywords and Boolean operators. Specifically, the first set contained the keywords that referred to serious language games, the second set involved the keywords concerning language learning and teaching, and the final set included the keyword referring to the digital version of the game. In each search run, the keywords were used in various combinations, including ("educational game" OR "serious game" OR "simulation game" OR "newsgame" OR "roleplaying game" OR "RPG" OR "pervasive game") AND ("language learning" OR "language teaching") AND ("digital"). The publications identified at this phase were pasted in an online document. The initial article search produced a total of n = 163 papers (records).

Each paper (record) was screened separately by two researchers. The reviewers worked independently while screening an assigned pool of records, eliminating those that did not meet the eligibility requirements. In case of disagreement, pairs of reviewers met to make the final decision concerning paper in/exclusion. Whenever needed, the third researcher was asked to help reach a consensus. Upon completion of this phase, n = 50 eligible studies published prior to April 2023 were selected for analysis.

To obtain data for analysis, an online data collection form was used in the present study. The form contained authors' names, years of publication, methodology (research design, sample, data collection), pedagogical characteristics (language, targeted language learning area, learning theory, pedagogical approach), game design characteristics (genre of the designed game, game design framework/principles, game elements, technology), and outcomes (language learning gains, non-language learning gains, emotional outcomes, behavioural outcomes, other). Each investigator was involved in data extraction from an assigned set of included papers. Next, four investigators reviewed, completed, and analysed the dataset, each considering one research question. The results of the analysis are presented in the form of frequencies and identified themes. The reviewed papers and full results are presented in the Appendices, available at https://rb.gy/mdkzf3.

4. Results and Discussion

4.1. Methodologies employed in the included studies

Various research designs were adopted and reported in the papers reviewed (Appendix 2). The most common methodological orientation was an ex-

perimental research design (n = 15), followed by a quasi-experimental design (n = 6). Much less frequent were studies adopting a research and development design (R&D) (n = 3), mixed methods design (n = 3), pre-experimental (n = 2) or case study designs (n = 2). Individual studies reported the use of design and development research, research and development design in combination with experimental design, a one group pre-test post-test research design, or user and field study designs. In the case of fourteen papers, the research design was not explicitly or clearly stated by the authors. These findings reflect those of Borona et al. (2018) and Almaki et al. (2023) who identified the quantitative approach as the dominant methodological choice for studies of 3D multi-user virtual environments in computer-assisted language learning and simulation games in K-12 education, respectively. However, it is important to note that the outcome in the present study differs from the findings of Acquah and Katz (2020) and Hung et al. (2018), who observed that the majority of research in DGBLL mainly favoured a mixed methods approach and did not report the use of R&D designs.

Considering the type of participant, as presented in Appendix 3, the bulk of the studies of serious digital language games reviewed involved learners in higher education (n = 18) and primary school (n = 15). Fewer studies were conducted among learners in secondary school (n = 2) or kindergarten (n = 1), and one study had participants aged 7–65 years old. Additionally, there were studies that included different types of participants, such as preschool and primary school learners, primary and secondary school learners. This is in line with the findings of Hung et al. (2018), but it differs from the results of Acquah and Katz (2020), who focused exclusively on participants of 6–18 years..

As presented in Appendix 4, the research reported in the reviewed papers included participants from several countries. Studies involving participants from Asia were most common (n = 28) and included seven papers reporting participants from Indonesia, with the same number from Taiwan, three studies from Iran, and two studies from China, Japan, Malaysia, and Turkey respectively. Single studies including participants from Hong Kong, Jordan and South Korea were also found in the review process. Studies with participants from Europe were far less frequent (n = 7), three of which had participants from the Netherlands, two from Greece and two from Spain. Concerning the participants from North America, three studies in the present review included participants from the USA. In one study participants came from South America, Brazil, and in one they came from Africa, Egypt. In addition, one of the reviewed papers involved participants from many countries. However, in nine papers, the authors did not (clearly) provide information about the country their participants came from. The observation concerning the

domination of studies conducted within an Asian context is in accord with Acquah and Katz's (2020) finding, which showed a substantial input from researchers conducting game-based language learning projects in East Asia and the Middle East.

Regarding the size of the samples in the reviewed studies, the number of participants ranged from 1 to 9 participants (n = 3), 10–49 (n = 19), 50–99 (n = 12), 100–999 (n = 12), and there were two studies including over 1000 participants. In the case of two papers, information about the sample size was unclear, or not provided (Appendix 5). Given the adopted research inquiry strategies described above, this diversity was not unexpected. Compared with previously reported results in simulation game studies in K-12 education (Almaki et al., 2023) and 3D multi-user virtual environments in computer-assisted language learning (Borona et al., 2018), the outcome of this study shows a greater diversity and larger sample sizes in the area of serious digital language games.

As presented in Appendix 6, the analysed studies gathered research data using various methods, in different combinations. The use of two types of data collection methods was most common (n = 18), followed by one type of data collection method (n = 15) and three types of data collection methods (n = 11). There were also papers that made use of four types of data collection methods (n = 5). In one study, the data collection method was not explicitly mentioned. Similar to the reviews conducted by Borona et al. (2018) and Almaki et al.(2023), this study also identified an array of data collection tools, including pre- and post-tests, questionnaires, surveys, observations, logs, journals, and more. However, what sets this study apart is its exploration of tool triangulation in serious digital language game studies, providing more insight into the reliability of the results of the papers reviewed.

4.2. Pedagogical characteristics of serious digital language learning games

Concerning the frequency of languages designed or used for such games, English is the most common language, appearing in forty-two studies. Spanish (n = 3) came next with two occurrences, followed by Chinese (n = 2), Japanese (n = 2), German (n = 2), French (n = 2), as well as Arabic and Italian, each n = 1. The numbers do not add up the total of 50, as several studies investigated more than one language. The choice of the target language does not differ from the choice reported in previous DGBLL review studies (Acquah, Katz, 2020; Dhimolea et al., 2022; Hung et al., 2018).

As to the language learning areas targeted in the games reported in papers reviewed, as shown in Appendix 7, the largest number of papers concentrated on vocabulary (n = 29), among which one study focused on vocabulary assessment (not teaching) and four papers implied vocabulary as a targeted area in the game. There are ten papers addressing the area of reading (including two papers implying this area and one assessing it), six papers addressed the area of speaking, and six papers investigated writing (again, one paper implied this area and one assessed it). In addition, five papers explored pronunciation (including two that implied this area) and five, listening. Finally, two papers addressed grammar, and one paper focused on spelling. Some research explicitly elucidated the scope of vocabulary and grammar targeted in the serious games. Among the few studies, Tlili et al. (2021) provided the names of animals, Khatoony (2019) of sports, animals, fruits, jobs and colours, while Alexandre et al. (2023) indicated grammar areas (pronouns, prepositions, determiners) as well as various categories of vocabulary (objects, clothing, professions and occupations, animals, food, family, colours, geometric shapes, human body). As in the above-mentioned systematic reviews, a variety of language learning areas were found to be addressed, including vocabulary, speaking, listening, reading, pronunciation, grammar, writing (Dhimolea et al., 2022; Hung et al., 2018; Weng, Chiu, 2023; Zhai, Wibowo, 2023), as well as spelling, phonological awareness, orthographic processing (Acquah, Katz, 2020), or cultural learning (Borona et al., 2018). However, our results differ from those by Almaki et al. (2023), who included serious games applied for teaching subjects such as mathematics, social science, or technology in their systematic review.

As presented in Appendix 8, an array of theoretical foundations was employed within the examined papers, despite the fact that many of the articles (n = 19) lacked explicit reference to any learning theories. Among the various theoretical perspectives, constructivism (n = 4) and cognitivism (n = 3) emerged as the most frequently mentioned frameworks. Social Cognitive Theory (n = 4), situated/contextual learning (n = 2), and the ARCS model (n = 5) also gained attention in the reviewed papers. Other frameworks were also used in individual studies. Overall, the theoretical foundations highlighted the significance of incorporating social interactions, real-world contexts, input and motivational factors in using serious games for language learning. This aspect was not explored in earlier DGBLL review studies. Our findings illustrate that the theoretical foundations seem strongly underdeveloped in the reviewed papers and neglected in other systematic reviews (Borona, 2018; Dhimolea et al., 2022; Naul, Liu, 2020; Weng, Chiu, 2023; Zhai, Wibowo, 2023). However, the present results are consistent with those obtained by Krath et al. (2021) in a systematic meta-review of theoretical foundations

in game-based learning research, who found that self-determination theory, flow theory, experiential learning theory, and constructivist learning theory were most common.

Various pedagogical approaches were employed in the selected papers (Appendix 9). Game-based learning (GBL) was the most frequently employed approach, with ten papers directly referring to GBL, seventeen to DGBL, and a few to video, mobile, AR, VR or AI game-based learning. Technology assisted language learning (TALL) was also a popular approach with nine papers on MALL and three papers on CALL. Other pedagogical approaches such as Content and Language Integrated Learning (CLIL), inquiry-based learning, online learning, autonomous learning, or other instructional strategies were also represented in the papers reviewed. The results highlight the wide range of pedagogical approaches adopted in the context of serious games for language learning, demonstrating the diverse strategies employed to enhance language acquisition through game-based instruction. Again, this is a novel outcome, not addressed in earlier DGBLL review studies.

4.3. Design characteristics of serious digital language learning games

Appendix 10 presents the explicitly stated genres of the games in the papers reviewed. Some papers (n = 5) which mention a specific genre declare it to be a role-playing game (RPG). Other explicitly stated genres are adventure games (n = 3), virtual games (n = 3) and visual novels (n = 3); murder-mystery and simulation games are each mentioned in two articles. Storytelling, action, survival, strategy, scenario and detective games are examples of the genres appearing in individual papers. In two articles, general information can be found describing the genre of the game used as a video game and a computer game, but no details are provided. Many of the games (n = 24) are not categorised with reference to a specific genre. Importantly, a number of games presented in the papers emerge as gamified tasks: from their description it can be learnt that they are digital versions of lexical activities used in coursebooks, such as wordsearch (Butler et al., 2014; Goumas et al., 2020), hangman (Logothetis et al., 2021), or jigsaw (Hwang et al., 2016; Sandberg et al., 2011). These results echo those obtained in Acquah and Katz's study, who found educational and mini-games, role-play, adventure, simulation and strategy game genres implemented in the studies, and in the Borona et al. (2018) study, which discusses the activities qualified by the researchers as "educational games" which are not games. However, the present study did not identify some of the genres found by Hung et al. (2018), such as music games, or alternate reality games.

There are a few mentions of frameworks or principles in the papers which directly relate to the design of the presented games. In the article by Hu et al. (2022) there is a game design framework, called an "AR educational game design framework" (p. 4). Mulder et al. (2021) report applying the Elo system for chess rating to provide each participant with a good amount of challenge. There are papers that allude to design principles or frameworks, such as Goumas et al. (2020) who discuss the design of their game; yet, in doing so, it appears that the researchers present the research procedures in their study but do not provide the specific design principles that guided the design of the game. Similarly, Lee (2022) mentions the ADDIE model; however, it seems that the model is used to describe an AR-based study, without specifically addressing game design principles.

Appendix 11 overviews game elements mentioned in the papers. As can be seen, the game element which was used most often in the research is story/narrative (n = 23). Levels (n = 15) and points/scores (n = 12) were also popular in creating serious language games. Social interactions and rewards are discussed in eight and seven papers respectively, while avatars (n = 6), teammates (n = 5) and leaderboards (n = 2) were less popular and used in fewer studies than other game elements. In contrast with the Acquah and Katz's (2020) study, in the papers reviewed there is no mention of 'competition' as an important game element, which may be a result of inconsistency in the perception of what should be and should not be considered good game elements. It is important to note that the total number of mentions is higher than the number of the articles reviewed as some of the papers report applying more than one game element. Appendix 12 displays the number of game elements used in the games. As can be seen, many papers (n = 12) either did not mention any game elements, or used only one (n = 12). Ten studies mentioned two game elements in their papers, while eight mentioned using three. Four elements were used in three studies, while five elements were implemented in five games reported on in the reviewed papers. It is worth noting that, although some papers do not explicitly mention game elements, they still describe them, e.g. when it is stated what the player needs to do to complete one level, or gain a reward/point, similarly to the ones discussed in Acquah and Katz's (2020) study. What is more, in the analysed publications, the terms "challenges," "missions," and "quests" are used interchangeably (Faizal, 2016; Ghani et al., 2019; Tlili et al., 2021). This is also the case for the terms "multiple players" and "teammates" (Butler et al., 2014), as well as "characters" and "avatars" (Lee, 2022). However, the studies which mention "leaderboards" state that they are visible to the students, teachers and researchers, but do not provide information about whether the students can see each other's results. These

are also novel findings, not reported earlier in previous DGBLL reviews, revealing current practices in the serious digital language games design, which are, apparently, not guided by game design frameworks or principles.

The papers reviewed report various uses of technology, classed here as hardware and software uses (Appendix 13). There are nine mentions of different pieces of hardware employed in the research. Most often used are computers (n = 3) and tablets (n = 2), whereas mobile phones, headphones, VR devices and other pieces of equipment are mentioned in one paper each. The software used to develop or adapt the games is named fifty times in various articles. The Unity game-development platform is mentioned in seven papers, and mobile applications/systems, as well as augmented reality, are reported in six articles each. Five papers state that some software was used, but do not specify which exactly. The use of the Android system is reported in four papers, while the categories of online games as well as virtual reality (VR) are named in three papers each, without specific systems mentioned. Artificial intelligence (AI), an online/offline game platform and Java (programming language) are each mentioned in two papers. Ren'Py (a visual novel engine), mixed reality (MR), web server, QR codes, Google Earth, Windows 7, adaptive digital learning environment, game software, open-data kit system and Adobe Systems are reported in one paper each. There are also eight papers in which it is not stated what technology was employed in the research. Similarly to previous reviews (Acquah, Katz, 2020; Hung et al., 2018), this study also found that computers and mobile devices were used for game delivery. Yet, unlike Hung et al. (2018), this review did not identify the use of video consoles, suggesting that existing serious digital language games have not yet been made available as digital video games, similar in quality to commercial video games designed for entertainment purposes, indicating a gap in the research and practice. With regard to the findings concerning software employed, these results augment our understanding of design practices in the field by providing insight into software choices made by serious digital language game designers.

4.4. Outcomes in serious digital language learning game studies

Five key outcome categories in serious digital language game studies were identified: language learning gains, non-language learning gains, as well as emotional, cognitive, and behavioural outcomes, broken down into positive and negative aspects.

In the category of language learning gains, nine domains were identified (Appendix 14). Many papers addressed the issue of vocabulary

(n = 21). Researchers also reported gains in skills of reading (n = 6), speaking (n = 5), writing (n = 4), pronunciation (n = 4), listening (n = 2), grammar (n = 1), and spelling (n = 1). Additionally, one article reported gains in the area of Chinese language arts. The current review corroborates findings of previous systematic reviews. Specifically, vocabulary gains are more frequently reported than other language areas, aligning with findings from research on digital DGBLL by Hung et al. (2018), VR games by Dhimolea et al. (2022), AI applications by Zhai and Wibowo (2023), and 3D multi-user virtual environments by Borona et al. (2018).

Non-language learning gains was the second category identified (Appendix 15). Most papers addressed the issue of collaboration / cooperation / teamwork / social interaction (n = 8). Additionally, two papers analysed signs of imagination / new ideas / creativity (n = 2). Self-reflection, confidence, relevance and geographical knowledge were each analysed by a different author. Among the fifty articles reviewed, only a few addressed non-language outcomes like collaboration and social interaction in online learning environments, with terms such as collaboration, cooperation, and teamwork often used interchangeably. Collaboration was the only outcome found, consistent with findings from Borona et al. (2018), Acquah and Katz (2020), and Hung et al. (2018), while other competencies they reported, such as critical thinking, creative thinking, and problem-solving were not observed in the current review. Content learning results were similar to those of Hung et al. (2018), who noted learner improvements in mathematics and culture, and Borona et al. (2018), who found increased cultural awareness.

The third category concerned emotional outcomes (Appendix 16), broken down into positive and negative affective factors. Most authors analysed positive affective factors, such as motivation (n = 17), interest (n = 10), fun (n = 7), engagement (n = 3), enjoyment (n = 4), satisfaction (n = 2), eagerness (n = 1) and enthusiasm (n = 1). Negative factors, such as anxiety (n = 2), frustration (n = 1) or nervousness (n = 1) were rarely analysed. The findings from the current review align with those of Almaki et al. (2023), who found evidence supporting gains in engagement and motivation in K12 learners through simulation games, and Zhai and Wibowo (2023), who reported enhanced learner engagement via AI integration. However, comparisons with previous DGBLL research are challenging due to varying categorizations of affective outcomes. Acquah and Katz (2020) grouped results into broad affective/psychological states, while Hung et al. (2018) identified specific affective outcomes such as positive attitudes and motivation. The present study not only confirms previously identified emotional outcomes such as positive attitudes, motivation, and engagement but also distinguishes a broader and more precise range of emotional states, separating them from cognitive outcomes.

Within the category of positive cognitive outcomes (Appendix 17), we identified learning strategies (n = 3), challenge (n = 4), attention / concentration (n = 4), reading strategies (n = 2), academic resilience (n = 1), and self-efficacy (n = 1). The current review acknowledges self-efficacy as a cognitive outcome, also reported in prior DGBLL reviews by Acquah and Katz (2020) and Hung et al. (2018). However, it uncovers a broader spectrum of cognitive outcomes within serious digital language learning games, while noting an absence of negative cognitive outcomes.

Behavioral outcomes within digital serious language learning games appear to be scant (Appendix 18). Only one paper (n = 1) identified by this review dealt with active participation. Lee's (2022) study, using an engagement questionnaire, suggests that the medium (print vs. AR technology) does not significantly impact students' behavioral engagement, provided that the teaching is designed following game-based and problem-based learning principles. This finding aligns with Acquah and Katz (2020) and Hung et al. (2018), who related participatory behaviours to various factors such as time on task and learner interactions. Nevertheless, the current review highlights a gap in the literature concerning the exploration of behavioural outcomes in this field.

5. General discussion and concluding remarks

Following PRISMA guidelines, this study conducted a systematic review of research on serious digital language learning games with two aims. First, this study aimed to map out the issues pertaining to S-DGBLL that are currently being addressed in empirical studies in the DGBLL area. The second aim was to use the findings to inform the development of a research-based framework of S-DGBLL expertise for L2 teachers. Fifty articles were reviewed, examining research methodologies, pedagogical characteristics, design, and outcomes in serious digital language usage learning games. The results show that the studies reviewed were predominantly conducted in the Asian context, mainly among students in higher education and primary school learners, on samples ranging from very small–up to five participants–to those with over 1,000 study participants. Serious digital language learning game studies tend to adopt either experimental (including quasi-experimental and pre-experimental) designs or designs associated with R&D, and employ up to four data collection methods. Next, the analysis revealed that a great majority of the papers targeted the English language and many concentrated on the design of games to support the development of vocabulary. Twenty-five learning theories and forty-three pedagogical approaches were identified in the papers re-

viewed. Further, the study showed the number and types of game elements incorporated, hardware and software used. It found a relative lack of game design principles or frameworks employed to inform the game design process. Lastly, four types of outcomes were found, i.e. language and non-language gains, as well as emotional, cognitive, and behavioural outcomes.

Contrary to our expectations, the systematic review did not provide a strong basis for identifying the competences required for language teachers aiming to utilise S-DGBLL in their classrooms. Instead, the review revealed four areas of expertise that appear to be essential for language teachers intending to integrate this approach for educational purposes. These include: (1) pedagogical competences, (2) game-design competences, (3) technology-related competences, and (4) research competences. The first component, pedagogical competences, form an essential part of the teacher S-DGBLL professional repertoire. This consists of knowledge of language learning theories (e.g., constructivism, cognitivism, dual-coding theory, interactionist theory), teaching methods (e.g., content and language integrated learning, data-driven learning, active learning, collaborative learning) and techniques (e.g., role-play, storytelling, simulation)–which can pave the way for serious digital language learning game design. The second area, game design competences, pertain to game design skills to be deployed in planning, designing and implementing games in L2 instructional contexts. The next component, technology competences, concern various game-related CALL competences which enable planning the whole process from the ICT perspective, including specific technological solutions such as VR, Augmented Reality (AR), and AI. Lastly, research competences include the skills associated with research design (e.g., quasi/pre/experimental, case study, R&D), data collection and analysis, which allow teachers to evaluate the games they design or choose for their learners, including the effect of games on (intended) learning outcomes. As such, our work advances the study on teacher roles in game-based learning (Molin, 2017) by focusing on the integration of serious games in L2 learning and identifying areas of L2 teacher expertise.

Before concluding, it is essential to acknowledge two main limitations in this study. First, when extracting and analysing data each researcher worked independently with a view to responding to a specific research question. Without an additional reviewer to check the extracted data, the results might be biased. In addition, this systematic review concentrated exclusively on English-language publications. This means that there might exist additional papers written in other languages that could have been overlooked.

Despite the limitations, this study enriches the literature in the field of serious digital games for language learning in two ways. First, it captures different aspects of S-DGBLL, providing an up-to-date overview of existing

issues addressed in the field. Furthermore, this research advances the litera-ture by identifying possible L2 teacher competences in the area of digital se-rious language learning game design and implementation. Even though the present study offers a preliminary framework, it makes emerging areas of expertise explicit, helping define strategies for teacher training.

Several avenues for future research can be offered. Firstly, there is a need to elaborate on the framework of teacher competence. This is why a further review–this time of publications specifically devoted to S-DGBLL pedagogies and game design–is needed. In addition to this, researchers are also encouraged to investigate game design principles and frameworks in the development of serious digital language learning games. Finally, emerging technologies, such as virtual reality, augmented reality, and artificial intel-ligence constitute a potential area for exploration for researchers interested in S-DGBLL.

The results of the study carry implications for practice within the area of L2 education. The increase in the number of the studies implies that there is a demand for serious digital games for language learning. Thus, teachers could use such games or create their own. What is more, as most of the games were used to teach English, teachers of other languages can analyse the strong and weak features of existing serious games and adapt them to the teaching other languages. Another suggestion for teachers is to ana-lyse the section describing the outcomes as all of these could be important in course-planning. Beyond this, teachers can develop and apply the com-petences identified (i.e., pedagogical, game design, technology, and re-search) into their teaching practice to provide more effective and engaging L2 instruction. Lastly, these competences can serve as a basis for the devel-opment of L2 teacher training programmes. With the framework provided, teacher educators can be better equipped to prepare teachers for the use of S-DGBLL.

Bibliography

Acquah E.O., Katz H.T. (2020), *Digital game-based L2 learning outcomes for primary through high-school students: A systematic literature review.* "Computers & Education", No 143, 103667. doi: 10.1016/j.compedu.2019.103667

Alexandre I.M., Lopes P.F., Borges C. (2023), *Roadmap for the Development of En-Lang4All: A Video Game for Learning English*. "Multimodal Technologies and Interaction", No 7(2), p. 17.

Almaki S.H. et al. (2023), *A systematic review of the use of simulation games in K-12 education*. "Interactive Learning Environments", No 5, pp. 1–25.

Blume C. (2020), *Games people (don't) play: An analysis of pre-service EFL teachers' behaviors and beliefs regarding digital game-based language learning*. "Computer Assisted Language Learning", No 33, pp. 109–132.

Borona S., Tambouris E., Tarabanis K. (2018), *The use of 3D multi-user virtual environments in computer assisted second language learning: a systematic literature review.* "International Journal of Learning Technology", No 13, pp. 249–274.

Butler Y.G., Someya Y., Fukuhara E. (2014), *Online games for young learners' foreign language learning*. "ELT Journal", No 68, pp. 265–275.

Chen H.-J.H., Hsu H.-L. (2020), *The impact of a serious game on vocabulary and content learning.* "Computer Assisted Language Learning", No 33, pp. 811–832.

Dahalan F., Alias N., Shaharom M.S.N. (2023), *Gamification and game based learning for vocational education and training: A systematic literature review.* "Education and Information Technologies", No 29, pp. 1279–1317.

Dhimolea T.K., Kaplan-Rakowski R., Lin L. (2022), *A systematic review of research on high-immersion virtual reality for language learning*, "TechTrends", No 66, pp. 810–824.

Faizal M.A. (2016), *The effects of conversation-gambits visual-novel game on students' English achievement and motivation*, (in:) 2016 International Electronics Symposium (IES), Denpasar, Indonesia, pp. 481-486.

Ghani M.T.A. et al. (2019), *E-Mufradat: a digital game application for learning Arabic vocabulary among non-native Arabic speakers*. "International Journal of Academic Research in Progressive Education and Development", No 8, pp. 58–69.

Goumas, S. et al. (2020), *Wordsearch, an Educational Game in Language Learning.* "Journal of Engineering Science & Technology Review", No 13, pp. 50-56.

Hu L. et al. (2022), *The Practice and Application of AR Games to Assist Children's English Pronunciation Teaching*. "Occupational Therapy International", No 2022, 3966740.

Hung H.-T. et al. (2018), *A scoping review of research on digital game-based language learning.* "Computers & Education", No 126, pp. 89–104.

Hwang W.-Y. et al. (2016), *Evaluating listening and speaking skills in a mobile game-based learning environment with situational contexts.* "Computer Assisted Language Learning", No 29, pp. 639–657.

Krath J., Schürmann L., von Korflesch H.F.O. (2021), R*evealing the theoretical basis of gamification: A systematic review and analysis of theory in research on gamification, serious games and game-based learning*, "Computers in Human Behavior", No 125, 106963. doi: 10.1016/j.chb.2021.106963

Lai K.-W.K., Chen H.-J.H. (2023), *A comparative study on the effects of a VR and PC visual novel game on vocabulary learning*. "Computer Assisted Language Learning", No 36, pp. 312-345.

Lee J. (2022), *Problem-based gaming via an augmented reality mobile game and a printed game in foreign language education*. "Education and Information Technologies", No 27(1), pp. 743–771.

Logothetis I. et al. (2021), *Transforming classic learning games with the use of AR: The Case of the Word Hangman Game*, (in:) International Conference on Human-Computer Interaction. Cham: Springer, pp. 47–64.

Maugard N. (2019), *Fundamentals of Serious Games,* (in:) Akhgar B. (ed.) Serious Games for Enhancing Law Enforcement Agencies. Security Informatics and Law Enforcement. Cham: Springer, pp. 15–41.

Mulder E. et al. (2021), *Serious game-based word-to-text integration intervention effects in English as a second language*. "Contemporary Educational Psychology", No 65, 101972. https://doi.org/10.1016/j.cedpsych.2021.101972

Naul E., Liu M. (2020), *Why story matters: A review of narrative in serious games*. "Journal of Educational Computing Research", No 58, pp. 687–707.

Page M.J. et al. (2021), *PRISMA 2020 explanation and elaboration: Updated guidance and exemplars for reporting systematic reviews*. "BMJ", No 372, n71.

Peterson M., Jabbari N. (2023), *Digital games in language learning: Case studies and applications*. Oxon, New York: Routledge.

Peterson M. et al. (2020), *A review of research on the application of digital games in foreign language education*, (in:) Kruk M., Peterson M. (eds.), New technological applications for foreign and second language learning and teaching. Hershey, PA: IGI Global, pp. 69–92.

Reinders H., Wattana S. (2015), *Affect and willingness to communicate in digital game-based learning*. "ReCALL", No 27, pp. 38–57.

Reinhardt J. (2019), *Gameful Second and Foreign Language Teaching and Learning: Theory, Research, and Practice*. Cham: Springer.

Sandberg J., Maris M., De Geus K. (2011), *Mobile English learning: An evidence-based study with fifth graders.* "Computers & Education", No 57, pp. 1334–1347.

Tlili A. et al. (2021), *A smart collaborative educational game with learning analytics to support English vocabulary teaching*. "International Journal of Interactive Multimedia and Artificial Intelligence", No 6, pp. 215–224.

Weng X., Chiu T.K. (2023), *Instructional design and learning outcomes of intelligent computer assisted language learning: Systematic review in the field.* "Computers & Education: Artificial Intelligence", No 4, 100117. doi: 10.1016/j.caeai.2022.100117

Zhai C., Wibowo S. (2023), *A systematic review on artificial intelligence dialogue systems for enhancing English as foreign language students' interactional competence in the university.* "Computers & Education: Artificial Intelligence", No 4, 100134. doi: 10.1016/j.caeai.2023.100134