



INSTITUTIONEN FÖR  
TILLÄMPAD IT

# VIDEO GAME-BASED LANGUAGE LEARNING AND LITERACY

A systematic review of the current state of research  
on video game-based learning and language  
learning

**Oscar Holst**

---

Thesis:	15 hp
Program:	Learning, communication and IT
Level:	Advanced
Year:	2022
Supervisor:	Sofia Serholt, Nataliya Berbyuk Lindström
Examiner:	Marisa Ponti
Report No:	2022:081

# Abstract

Over the last decades, the video game medium has become increasingly popularized and technologically advanced. Consequently, the learning sciences have started to investigate the pedagogical potential of this medium. Literacy researchers have argued that unique features of the video game medium offer possibilities for effective language learning.

This study sets out to identify what theories guide games design utilized for language learning, what game designs were consequently considered relevant, and what learning gains were identified using these games. To answer this question, an *inductive systematic literature review with a qualitative approach* was conducted.

The study relates to the fields of *game-based learning*, *literacy* and *language learning* and these concepts informed the search strategy. The search was carried out in five databases, and 21 articles were ultimately included in the study. Using *constant comparative analysis*, four categories of studies were identified and underlying theory, design, and learning gains were described.

The results suggest that video games designed using theories from the fields of literacy or cognitive attention can be used to give a boost to fundamental literacy skills such as phonological awareness, but also in some cases vocabulary, among children around the age of school entry. Studies on serious games targeting primarily vocabulary were difficult to analyze since theoretical and methodological approaches were very diverse. Studies investigating serious games centered around a narrative report problematic results. Case studies on commercial games identify language learning opportunities using video games, but struggle with the question of generalizability and applicability.

Theoretically, the review points out that the concept of game-based learning is not entirely easily defined, since the concept of “game” is sometimes used in tandem with other concepts which highlight different aspects of the design. Additionally, it is difficult to tie specific game design elements to successful learning outcomes. Practically, the review suggests that game-based learning may be of particular value to disadvantaged schools or individuals in order to promote fundamental literacy skills.

## Key words

game-based learning, language learning, literacy



# Table of contents

1	Introduction .....	1
1.1	Goal and aim .....	2
1.2	Research questions .....	2
2	Previous research.....	3
3	Method.....	6
3.1	Methodology: Literature reviews .....	6
3.2	Procedure.....	7
3.2.1	Define.....	7
3.2.2	Search.....	9
3.2.3	Select – criteria of inclusion/exclusion .....	10
3.2.4	Analysis .....	11
3.2.5	Presentation.....	11
3.3	Ethics.....	12
3.4	Validity and problems related to method .....	12
4	Results .....	14
4.1	THEORIES USED IN RESEARCH ON VIDEO GAMES AND LANGUAGE LEARNING .....	15
4.2	DESIGN OF VIDEO GAMES USED FOR LANGUAGE LEARNING.	21
4.3	LEARNING GAINS .....	28
5	Discussion.....	37
5.1	Comparisons with previous research .....	37
5.2	Theoretical contributions .....	39
5.3	Practical implications .....	41
5.4	Suggestions for future research .....	42
5.5	Limitations .....	43
6	Conclusion.....	45
7	References .....	46
8	Appendix .....	59
8.1	Glossary .....	59
8.2	Overview of articles .....	61



# 1 Introduction

Learning and school results of children and young adults are intimately tied to their language development. For instance, a child's early reading proficiency is a strong predictor of later academic achievement (see for instance McLoughlin et al., 2014), but may also have implications for phenomena such as physical and mental health (DeWalt et al., 2004). However, literacy research has pointed out that with social, cultural and technical changes in society, practices around reading and learning also change (Mills, 2015). The advancement of technologies such as e-books and video games forces the question of what opportunities, as well as potential problems, that this development might entail.

According to the UK Association of Interactive Entertainment (2022) the video games industry is still rapidly growing with a year-on-year increase of revenue over the last couple of years. Additionally, gaming is now a widely used medium across groups. 50% of people aged 6-64 within the European key markets (France, Germany, Italy, Spain and the UK) play videogames. 68% of 6–10-year-old children and 79% of 11–14-year-old children do the same. Additionally, 47% of European video gamers are now women (UK Association of Interactive Entertainment, 2022). One of the imperatives behind investigating game-based learning is the fact that games have become so widely available, no longer limited only to computers and gaming consoles but now also being able to be played on phones and tablets. Thus, they are widely used by children (Kim et al., 2017) This warrants the question both of the quality of existing, commercially available products but also the question of what researched products can bring to the table (Dore et al., 2019)

In tandem with this development, research on so called game-based learning has asked questions of how games can be used to foster learning generally. There are indications that games may lead to content learning, such as in mathematics (Kebritchi & Hirumi, 2008; Samur, 2012), science (Miller et al., 2011) and history (Huizenga et al., 2009). In the early 2000s, the New Literacy researchers called for increased attention to the connections between gaming and *language* learning more specifically, arguing that the new literacies provided by the gaming medium would provide news ways of learning language (Gee, 2003). On the other hand, other researchers have suggested that while game-based learning may seem promising, there are a number of problems associated with it (Linderoth, 2012).

Given the rapid development described above, research should endeavor to carefully explore the connection between language learning and video games. This

study thus sets out to investigate the question: what is the current state of empirical research around language learning utilizing video games?

## 1.1 Goal and aim

The aim of this study is to investigate the current state of research regarding the use of video games for language learning. More specifically, the goal has been to identify the theoretical perspectives guiding research on how video games can be utilized, what design elements are considered relevant in the video games, and what learning outcomes are achieved by using video games for language learning. Using this approach, lessons learned from previous studies may help to inform future research endeavors but may also be of interest to anyone interested in using, or even designing, games for language learning purposes.

This study is an *inductive systematic literature review with a qualitative approach* (Wolfswinkel et al., 2013). While individual studies present quantitative data that may be considered evidence, for the purposes of this essay they will thus be interpreted qualitatively. For the purpose of the question of effectiveness or evidence, the term *significant* will be used, when reporting on experimental studies with statistically significant quantitative results (as measured by ANOVA-analyses or similar measures).

## 1.2 Research questions

Three questions were formulated. I consider the answers to these questions to be related and have bearing on one another; theoretical underpinnings influence design decisions made when planning game-based activities, which in turn have different implications for the learning processes and the learning gains evaluated towards the end of the respective study.

The questions are:

1. What are the theoretical perspectives guiding video game based learning activities intended to enhance language learning?
2. Based on those theories, what game design elements were considered relevant when constructing or analyzing the games utilized in the selected studies?
3. What were the learning gains identified in the selected studies?

## 2 Previous research

A few previous literature reviews (all peer-reviewed), researching the relationship between game-based learning and language learning, were identified before carrying out this study. In this section, they are presented in the chronological order in which they were made. While there are many overlapping areas of interest between this study and those reviews, this study differentiates itself from them primarily in its focus on how theoretical considerations inform design and affect learning outcomes. Similarities and differences with regard to results will be brought up in the discussion section.

Lieberman et al. (2009) carried out a systematical literature review of studies on the effects of digital media (among them computer games) on children in the ages three to six. The included studies demonstrated effects on thinking, problem solving skills, language skills, mathematical skills, creative skills, collaborative learning, reading and planning skills. The part of the review which focused on language skills consisted of only a few studies, but these indicated that well-designed serious games could be effective for such areas as learning the alphabet, phonics, word recognition, spelling and writing, but also for learning a second language. On the other hand, Lieberman et al. (2009) point out that since so many commercially available products have never been researched on, it is difficult to know how digital products *generally* affect children's learning (Lieberman et al., 2009).

A meta study by Rosman et al. (2013) evaluated nine studies carried out between 2003-2011, focusing more specifically on the key terms "video games" and "learning". A primary focus of this meta study was the methods used and it was found that most of these studies were literature studies, followed by case studies and then experimental studies. Another observation made was that the studies employed a combination of quantitative and qualitative methods. Rosman et al. (2013) concluded that the studies, taken as a whole, showed that video games have potential as effective teaching aids and can motivate students to improve their language skills.

Hung et al. (2018) made a scoping review on the effectiveness of digital games in education. 50 articles were included, carried out between 2007 and 2016. In their analysis, the authors list eight particularly interesting themes found in the material: like Rosman et al. (2013) they observed that the most common methodological approach was mixed methods. The most commonly studied type of game was so called massively multiplayer online role-playing games (MMORPGs), but most of

the games were made or at least modified by the researchers. PC was the most common platform used as a gaming platform. Finally, most of the studies were focused on learning English as a second language or foreign language. Most of the studies had participants of differing language skills. University students were the most common type of participant. Finally, most studies drew the conclusion that the interventions had had positive effects on the learning of the participants, particularly with regard to their emotional or psychological status but also on their language learning. Hung et al. (2018) drew the conclusion that much of the research within DGBLL has focused on the acquisition of vocabulary, and for this reason they ask for studies that to a higher degree examine other aspects of language learning. They find it surprising that not more studies have investigated what they call contemporary competencies such as communicational- collaborative- or creative skills. They believe that games should be advantageous arenas for socialization and since there is a demand for research on how to foster 21st century skills they find it surprising that not more research targets this area. Finally Hung et al. (2018) ask for more reviews to be carried out in order to be able to identify additional missing pieces of this particular research puzzle (Hung et al., 2018).

Even more recently, Xu et al. (2020) carried out a scoping review in order to investigate research on digital game-based learning intended to support, more specifically, English language learning. Just like Hung et al. (2018) they drew the conclusion that most research focused on the acquisition of vocabulary. However, differently from them, they observed that mainly quantitative methods were employed in the selected studies. Another issue that was raised is that conclusions about learning must be somewhat cautiously analyzed. They believe that researchers have not been rigorous enough when it comes to reporting the language proficiency levels of their subjects. Another related problem is that when researchers use their custom-made games, the evaluation of skills may be problematic because the language challenges of the custom-made game may be poorly matched to the participants' levels of proficiency. For this reason, the writers advocate the use of commercial games to a higher degree since they draw the conclusion that the language of these games generally match the proficiency levels of the participants better. Additionally, they conclude that commercial games were considered to have "better" game elements (though what exactly is meant by this is not entirely clear to the reader). For this reason, the researchers recommend future studies to employ commercial games to a higher degree than previously (Xu et al. 2020).

This review differentiates itself from these prior reviews (apart from the fact that it incorporates more recent research) by employing the perspective that language learning consists of distinct dimensions such as phonology, morphology, semantics and pragmatics, rather than a view that skills (reading, writing, speaking and listening) or distinctions between first- and second language learning should be the only vantage points worth exploring. Compared to them, it also focuses on what theories

and designs were identified as relevant for the games, in addition to just reporting on the effectiveness of the games.

## 3 Method

This study is a systematic literature review. This method was chosen because it was believed that it would be difficult for any single study to properly answer the research questions. The purpose of this section is to present the methodological considerations that were made while preparing this study. Firstly, methodological perspectives on this method will be presented, in other words what theoretical perspectives that inform this chosen method (Cohen et al., 2018). The general purpose of making literature reviews will be brought up. Secondly, the procedure followed when carrying out the study will be presented. Lastly, some considerations will be made regarding the question of validity and ethics.

### 3.1 Methodology: Literature reviews

While reviews of research are always included in academical papers to some degree, the *systematic* review is considered its own method in that it is intended to produce new knowledge by compiling and analyzing all previously published research (Eriksson Barajas et al., 2013). The methods used for identifying this research are openly declared and they are thus open to scrutiny on a different level than literature reviews of other types of scientific studies. The purpose of a review can be for example to inform a certain practice, provide decision makers with critical information, or make suggestions for or outline future research endeavors. Reviews can contribute by clarifying and elucidating the current state of research. There are no general demands on how many studies must be included in a systematic literature review. Both quantitative and qualitative studies might be included in a single review, though different theorists have slightly different views on this (Eriksson Barajas et al., 2013).

Two principal guidelines have been used in order to inform the design of the method of this study. The first is the Campbell Corporations criteria for systematic literature studies (referenced in Eriksson Barajas et al., 2013). According to these criteria, a review must consist of:

1. Clearly described criteria for and methods for identification and inclusion of articles
2. An expressly described search strategy
3. Systematic coding of all included articles
4. Meta analysis to compile studies whenever possible (Eriksson Barajas et al., 2013, p 27)

Secondly, an inductive approach was decided upon. The method was modeled after Wolfswinkel et al.'s (2013) framework for what they *call inductive systematic literature review with a qualitative approach*. According to this model, the systematic literature study should be carried out in five steps. The first step is called *Define* which means that a focus for the study is identified and criteria for inclusion and exclusion are decided upon. The second step, *Search*, consists of carrying out trial searches (in order to establish that the search strategies seem to be functional), and then the actual searches are carried out in relevant databases. In the third step, *Select*, the hits generated by the databases are systematically assessed for relevance and relevant articles to be included in the study are identified (based on criteria of inclusion and exclusion). In the fourth step, *Analyze*, these articles are then read through in their entirety and relevant sections are marked. In the final step, *Present*, the report is written, and the results are presented. This procedure was followed while carrying out this study. The more specific considerations made during each step are presented in the next section.

## 3.2 Procedure

### 3.2.1 Define

In accordance with Wolfswinkel's et al. (2013) previously mentioned five-step method, the first step (*define*) of a systematic literature review is to identify the relevant field, or fields, of study in order to be able to translate the aim and the research questions into viable search strategies. Additionally, while this study is largely inductive in its approach, it was necessary to define the areas of study also for the purpose of making the study manageable. Given the aim, this study seeks to identify research articles that connect game-based learning and language learning and literacy research. The following is how these concepts were understood when defining the study:

#### *View of game-based learning*

Game based learning, in other words learning while playing different kinds of games, has been proposed to have a high degree of potential on the grounds that games can be very captivating to children and create engagement in a way that other modes of learning are less successful at (Gee, 2007; Prensky, 2001; Shute et al., 2009). Features of games that can contribute to engagement are control (the player is allowed to make choices in the game and observe the outcomes), clear goals, rewards, challenge, and feedback. Games can also create a sense of immersion for example by containing fantasy elements (Abdul Jabbar & Felicia, 2015; Garris et al., 2002; Ke & Abras, 2013; Lepper & Malone, 1987) In order to narrow this study down, it was decided to only include games that could be considered

”video games”, understood as games played using a screen such as a tv-monitor, a computer screen or the screen of a phone. Other forms of games were excluded.

### *View of literacy*

Literacy researchers such as Gee (2007) have argued that games foster other learning strategies than traditional media, indicating that conditions for language learning may be different than for example by reading a book. Literacy research is historically related to research on reading and language development, but the complexity of the concept is illustrated in Mills’ book (2015) where each chapter brings up a different branch of literacy research. According to Mills, there was a more universally accepted definition of literacy in the 1980s, but since then understandings of the concept have become more diversified. Generally, it can be said that while literacy researchers in the past had a more traditional view of what counted as reading (where a reading activity would usually involve some form of physical, printed text), the so called digital turn has made many literacy researchers reconsider this notion. Lines have blurred considerably between different forms of communication, where text or fragments of text is deeply embedded in and more or less prevalent in different forms of modern media (Mills, 2015). In this review, the concept of literacy is understood as being connected to language learning and game-based learning (possibly working as a sort of bridge between the other two fields). The term literacy is interpreted in the widest sense, meaning that the inclusion of articles is dependent on how the term is understood by researchers in the field.

### *View of language*

In this study, general language abilities are considered to consist of three dimensions: form (relating to phonology, morphology and grammar), content (semantics or vocabulary) and use (pragmatics) (Lundberg, 2010). These general abilities are in turn considered to be closely related to reading abilities and writing abilities. Reading abilities are considered to be related to decoding abilities (see the glossary for definitions of these concepts) as well as understanding, in other words being able to interpret, evaluate and possibly revise a written text. Writing abilities are considered to be related to the generation and organization of ideas and being able to choose words to represent these ideas (Elwér et al., 2013; Juel, 1988; Wengelin & Arfé, 2018) Additionally, when learning other languages other than our mother tongue, there may be particular implications. Therefore, second language acquisition was identified as a relevant area to include in the study. However, though it should be noted that there may be significant differences between learning a first and a second language, these considerations are discussed only as much as this problem is discussed in the articles in question.

After defining these fields as the relevant focus of this study the next step was to formulate a search strategy that could generate research connecting these fields.

### 3.2.2 Search

In the second and the third step of the process, *search* and *select* (Wolfswinkel et al., 2013), relevant articles are identified using a two-step process, first an automated step dependent on the algorithms of the databases, and then a manual one dependent on the evaluation by the researcher. Criteria of inclusion and exclusion decide which articles are considered relevant.

In the second step, *search*, (Wolfswinkel et al., 2013) the search strategy is planned and eventually decided upon. This is an iterative process in that test searches are initially carried out in order to provide an idea of whether the search strategy must be narrowed down or expanded. It must be manageable for the researcher to have time to go through all of the hits generated during the search, while still providing enough material to reach the goals of the study. By using the boolean operator AND the search can be narrowed down, and by using the boolean operator OR the search can be expanded. It is also considered important to include synonyms in order to cover research that may use slightly different terminology but at the same time largely studies the same phenomenon. Once the search strategy had finally been decided upon, the actual search was carried out and the number of hits was logged.

In collaboration with my supervisor, five databases were identified as relevant. These were ERIC, Proquest, Scopus, ACM and Web of Science. They were chosen on the grounds that they are well-known databases that (to varying degrees) contain articles that deal with connections between technology and learning.

The search strategy was kept constant across databases with only minor alterations when required to comply with the specific design of the database's search function. These variants are presented below.

<b>ERIC</b>	"Anywhere except full text"
<b>Proquest</b>	"Anywhere except full text - NOFT"
<b>Scopus</b>	Title, abstract, keywords
<b>ACM</b>	Title, abstract, author keyword
<b>Web of Science</b>	All fields

The search strain finally decided upon, and entered into each of these five databases, was the following:

**Literac\* AND ("video game\*" OR "digital game\*" OR "computer game\*" OR "tv game\*" OR "mobile game\*" OR "serious game\*" OR "commercial game\*") AND ("language learning" OR "language acquisition" OR "language development" OR "language skills" OR CALL OR "computer assisted**

**language learning" OR "communication skills" or "communication development" OR "vocabulary learning" OR "vocabulary acquisition" OR "vocabulary development" OR "pragmatism" OR "pragmatic development" OR "writing skills" OR "writing development" OR "reading skills" OR "reading development" OR phonetic\* OR "second language acquisition" OR "second language learning" OR DGBLL OR "digital game based language learning")**

After initial test searches, where no exclusions were made depending on articles' time of publication, the decision was made to only include articles published between 2017 and 2021. This was because the initial searches included too many hits to go through in a reasonable amount of time. But this decision also meant that the study focuses on only newer research, which was seen as an advantage. Previous literature reviews (see the section on previous research) only included research published up until 2017, so it was seen as an opportunity to include research from that year and forward in order to focus on more recent research in this review.

After articles from before 2017 were excluded, the final result consisted of 83 hits on Eric, 296 hits on Proquest, 86 hits on Scopus, 233 hits on ACM and 65 hits on Web of Science. It should be noted that these hits were accumulated in a logged in state using an account at Gothenburg University. Without being logged in to such an account, it is possible that some of these hits would not have been accessible. The final search was carried out on January 31, 2022. Finally, it should be added that given that the amount of material generated during this search was already considered to be quite extensive, it was decided that no further articles would be added even if they were published while carrying out the remainder of the study.

### **3.2.3 Select – criteria of inclusion/exclusion**

In accordance with Wolfswinkel's et al. (2013) suggestions, the third step of the process was to manually *select* relevant articles for inclusion and analysis.

The first step of this process was to read the titles and abstracts of all articles that had been included in the automatized search process. The articles that matched the following criteria of inclusion were singled out and included:

- the study needed to carry out some form of empirical research
- the study needed to explore the relationship between game-based learning and language development, in formal or informal learning settings
- the study needed to assess language learning outcomes of game-based learning
- the study needed to be a peer-reviewed research article (this was checked not only using the relevant database in question but also using Ulrichweb and the Norwegian list)

All articles that did not match all of the above criteria were excluded, but in addition to this:

- Whole books and conference papers were excluded
- Duplicates were identified and removed
- In a few cases, articles that would have passed this step based on the content of their title, keywords or abstract ultimately had to be excluded because they could not be accessed even with a GU account

A number of articles that may have been considered, but ultimately were excluded, were articles that did not have as their primary purpose to investigate the connection between game-based learning and language development.

After all articles had been manually gone through according to the abovementioned principles, 26 articles remained which would then, in the next step, be read in their entirety.

### **3.2.4 Analysis**

In the next step of the process, the articles which had passed the manual selection process based on an assessment of titles, keywords and abstract were then read in their entirety. In accordance with Wolfswinkel et al.'s (2013) suggestions, important passages relating to the research questions were identified and highlighted.

In the case of a few of the articles, after having read them in their entirety, it was ultimately decided to exclude these as well since the focus of these papers did not fully meet all the inclusion criteria described in the previous section. Ultimately, the final tally of included articles amounted to 21.

After reading all the articles, the analysis process started in earnest, using a grounded theory approach with no predetermined categories, but where constant comparative analysis (Wolfswinkel et al., 2013) was continuously employed as a strategy to identify patterns in the articles. Recurrent themes and topics then informed the presentation of the results and the discussion sections.

### **3.2.5 Presentation**

Barajas Eriksson et al. (2013) suggest that literature reviews incorporating both qualitative and quantitative studies may choose to present them separately from each other. The majority of the included studies, however, used a mixed-methods approach, employing both qualitative and quantitative research elements. For this reason, given the chosen inductive approach, a decision was made not to include or exclude studies based on whether the study could be classified mainly as a qualitative or quantitative study but rather on whether or not the study seemed to shed light on the research questions. It was also decided to present the articles using a thematic structure rather than one based around choice of method. For the sake of

identifying similarities and differences however, the studies were sorted in four categories.

### **3.3 Ethics**

Since a review does not carry out its own empirical study, the ethical perspectives mainly regard how the included studies themselves were carried out. While going through all the included articles, I have aspired to make sure they all had made ethical considerations, or had received ethical approval from an ethics committee. Though these aspects are described in more or less detail depending on the article, ethical considerations had been made to some degree in all included studies. For details regarding the ethical considerations of each individual study, I refer to the articles in question.

### **3.4 Validity and problems related to method**

The concept of validity regards whether the researchers are able to measure what is actually claimed (Barmark & Djurfeldt, 2015). Obviously, this would be a challenge in a review which seeks to synthesize results of studies that are difficult to compare in many ways. Regarding this aspect, it is important to note how the choice of research design correlates with potential claims of evidence.

With this in mind, the different studies all contribute to answering the research questions but in different ways and with different sorts of claims. The included studies could be classified as either experiments (with a degree of randomization between experimental group and control group), quasi-experimental studies (where randomization of participants in experimental group and control group has not been possible), and non-experimental studies (such as case studies) (Cohen et al., 2018). Within educational research, quasi-experimental designs are common since the randomization of participants into different groups is often not possible (Cohen et al., 2018), a problem which is brought up in number of included articles. Most of the studies in this review could be attributed to this category. Additionally, even though numbers are of interest for the study, qualitative dimensions were deemed necessary in order to answer the research questions. Therefore, this approach was chosen, though not being unproblematic.

Another aspect related to the validity of a review is whether the approach truly is systematic, and that the research process is transparent (Eriksson Barajas et al., 2013). The process followed is presented under the heading "Procedure". Additionally, I have aspired to represent the contents of the selected studies in as a correct manner as possible, which of course is a tremendous challenge when so little space can be offered to each individual study. Even with the best of intentions, there is always a risk of misinterpretations of the studies, or unintended skewing of results in

one way or another. The question of accurately representing the results of different studies is also connected to the problem of validity. Accurate descriptions and interpretations of the included articles and their results is necessary in order for the review to reach a desired level of validity (Eriksson Barajas et al., 2013).

Finally, an aspect related to the question of validity is to make sure only articles considered of high enough scientific quality are included in the review. This is indicated by the articles having been peer-reviewed (Eriksson Barajas et al., 2013). Only peer-reviewed articles were included in this review.

## 4 Results

Four major categories of articles were identified in the material. Note that the language aspect *vocabulary* is examined in some articles in the first category. The difference compared to the articles in category 2 lies in that the latter focused *primarily* on vocabulary.

Category	Theories	Design	Studied language aspects	References
<b>1:</b> Serious games targeting fundamental literacy skills (pre-reading abilities) in the first language	Mainly theories on literacy: The Simple View of Reading, visual-graphical theory, attentional control theory	Multi-choice: e.g. connect letters, syllables or words with auditory stimuli Moving objects/action elements	Phonological awareness Letter recognition Spelling Syllable recognition Whole word reading Reading fluency Vocabulary	Amorim et al. (2022) Baker et al. (2017) Bertoni et al. (2021) Pasqualotto et al. (2022) Ronimus et al. (2019) Samur (2019) Sánchez-Vincitore & Cross (2021) Schmitt et al. (2018) Serra & Gilabert (2021) Serret et al. (2017) Thomson et al. (2020)
<b>2:</b> Games primarily targeting second language vocabulary acquisition	Game theories such as the RETAIN-model	Multi-choice: e.g. connect words with pictures	Second language vocabulary acquisition Morphological awareness	Calvo-Ferrer (2020) Dore et al. (2019) Ensmann (2021) Rankin et al. (2021) Vasileiadou & Makrina (2017)
<b>3:</b> Serious games built around a narrative	Mainly literacy theories (visual literacy, argument literacy)	Narrative Multi-choice Exploration	Semantics Literacy (Visual/Argumentation) Writing	Huynh et al. (2021) Lawrence & Sherry (2021)
<b>4:</b> Inductive studies on commercial games	Mainly inductive approach	Mixed/complex Communicative elements	Pragmatics Semantics Spelling Morphology/grammar	Azman & Dollsaid (2018) Hobbs et al. (2020) Vasquez-Calvo (2018)

## 4.1 THEORIES USED IN RESEARCH ON VIDEO GAMES AND LANGUAGE LEARNING

The studies of the first category (serious games targeting fundamental literacy skills or pre-reading activities) use the following theories:

Within this category, theories from the field of literacy dominate, while theories related to game theory are only found in a few studies. Particularly, the Simple View of Reading (Gough & Tunmer, 1986) is highly influential, with a number of studies and games referring to it as a foundation for the game design. This theory points out that reading skills are dependent not only on developing conceptual understanding, but also on improving the technical skill called decoding. Decoding means being able to correctly break down whole words into their individual language sounds. Decoding is in turn considered dependent on phonological awareness, in other words being able understand the relationship between individual language sounds and their corresponding graphemes (letters).

Baker et al. (2017) point out that the theory of the Simple View of Reading (Gough & Tunmer, 1986) constitutes an important theoretical foundation (Adams, 1990; Ehri, 2005) for the design of GraphoGame, a game researched in a number of included studies. The concept of phonological (and more specifically phonemical) awareness is highlighted within this theory, but also generally pointed out as highly significant within several studies. For example, Amorim et al. (2022) base their work on meta-analyses of Ehri et al. (2001) who similarly highlighted the significance of phonemic awareness, and who also concluded that phonemic awareness programs *can* improve reading (Ehri et al., 2001). Ronimus et al. (2019) refers to a meta-analysis by Galuschka et al. (2014) whose results strongly support the phonological theory, claiming that only phonics instruction produces significant effects on reading performance for dyslexic individuals. Sánchez-Vincitore & Cross (2021) refers to research highlighting the importance of the alphabetic principle and developing phonological awareness (Ham et al., 2015, Ehri et al., 2001; Taylor et al., 2017).

A few studies point to the significance of the phonological theory with the addition of related sub-skills. For instance, Samur (2019) refers to the work of Lundberg et al. (1988) who point out that skills like letter recognition increase reading performance. Schmitt et al. (2018) point out the relevance of alphabetic knowledge, the mastery of uppercase and lowercase letters of the alphabet (Grant et al., 2012) along with phonemic awareness (Grant et al., 2012; Wagner et al., 1997)

The theory of the Simple View of Reading (Gough & Tunmer, 1986) also guided Serra & Gilabert's (2021) perspective on language, with Cain et al.'s (2015) addition to the theory that vocabulary with word recognition skills becomes more important after the acquisition of the alphabetic principle.

Similar to other studies of this category, Serret et al. (2017) bring up the Simple View of Reading (Gough & Tunmer, 1986) as a theoretical foundation for their work, and acknowledge the research highlighting the importance of phonological awareness. However, they argue that there is also research supporting the idea of advantages for visual-graphical, rather than phonological, learning for children with autism spectrum disorder, or ASD (Kamio & Toichi, 2000; Koshino et al., 2005). It has been suggested that visual decoding of whole words may be enhanced in children with ASD (Grigorenko et al., 2003), and that visual thinking may be a strategy to compensate for language limitations (Kana et al., 2006). Compared to other studies of this category, they thus consider the possibility that reading whole words as "pictures" rather than learning phonemes and graphemes individually, may still improve phonological awareness. Serret et al.'s (2017) interest in the concept of game-based learning comes from their understanding that many children with ASD are greatly interested in computers and technology, and theorize that this could be exploited to improve engagement and concentration also in academic activities (referring to Williams et al., 2002; Tuedor, 2006). Additionally, the authors argue that many children with ASD prefer the kind of organized, logical environments that games generally offer (Baron-Cohen, 2006). It was considerations like these, rather than any theory from the field of game-based learning, that the design of the game used in their study, SEMA-TIC (a game employing mainly visual-graphical strategies), was then based upon.

Two studies that also mainly target phonological awareness take a different theoretical approach. Bertoni et al. (2021) and Pasqualotto et al. (2022) take theories on attentional control, described in both of these articles as a more domain-general ability, as the starting point for their studies. However, there is still a strong link with the phonological paradigm described above in that these skills are considered early predictors of later reading skill, particularly for decoding (Franschescini et al., 2012). According to these theories, flexibly distributing attention of what is relevant for a task at hand and filtering out irrelevant distractions are skills needed for reading. For example, this is required when grouping letters into words and phrases or when moving from one line to the other when reading (Franceschini et al., 2012; Grainger et al., 2010; Grainger et al., 2016). In other words, in this paradigm it is the action elements commonly featured in games which is the element of interest and other theories from the field of game-based learning are not referenced. From a language perspective, these skills are also important when making close phonological, orthographic or semantic distinctions, in other words being able to make out small but crucial differences between language sounds, "images" of words, or

differences in meaning (Strijkers et al., 2015; Ruz & Nobre, 2008). Cognitive flexibility may also be relevant for making switches between different dimensions of writing, such as phonology, morphology, semantics and syntax, easier (Blair & Razza, 2007; Colé et al., 2007; Cartwright et al., 2017; Cartwright et al., 2019).

By contrast, though theories from the field of literacy were common in this category, theories from the field of game theory were very uncommonly referenced (other than previous research on games being presented as a background): Samur (2019) considered more general theoretical notions through reading of previous literature: he identifies nine game components central for good serious game design, divided into three groups: cognitive features (clear goals, immediate feedback, challenges and skills), affective features (elimination of the anxiety of failure, support for internalization and motivation, and behavioral features (articulation and reflection, robust reward system). These features are claimed to lead to engaging instruction (Atkinson & Hirumi, 2010; Bowman, 1982; Csikzentmihalyi, 1996; Gee, 2007; Gee, 2010a; Gee 2010b; Kiili, 2005; Prensky, 2001; Schell, 2008; Squire, 2003; Stahl, 2006). Taking all these elements into consideration, the ADDIE design model (Gustafson & Branch, 2002) was then utilized to design the game. The ADDIE model suggests that design be performed through a process of five steps: analysis, design, development, implementation and finally evaluation. Serra & Gilabert (2021) consider the notion of adaptability as the theoretically relevant aspect of game-based learning, indicating the ability of a system to adjust instruction based on learner abilities or preferences, and with the goal of improved learning (Oxman & Wong, 2014; Vandewaetere et al. 2011) The game-based theory guiding the design of the PBS KIDS Island, a website offering an array of game activities through a sort of hub, researched by Schmitt et al. (2018), was Fisch's capacity model (2016). This model directs attention to the highly limited cognitive resources of small children, indicating that when designing games for them, aspects such as controls need to be intuitive and logical. Fisch (2016) used the term *gameplay dominance* when the child focuses on gameplay issues unrelated to their learning. Additionally, this model suggests that games targeting this group should employ a "learn as you go" approach, avoiding lengthy tutorials (since the children are still developing their executive functions), repeating lessons across contexts, and be designed to be engaging, motivating and playful (Fisch, 2016). Also, Schmitt et al. (2018) consider modern interpretations of Vygotsky's (1930-1934 – 1978) sociocultural theory, claiming that games can serve as a tool to provide scaffolding through the zone of proximal development (McCarrick & Li, 2007; Wartella et al., 2016; Yelland & Masters, 2007). The game can do this for instance by providing feedback and hints if the child gets stuck, and provide less support as the child becomes more proficient (Yelland & Masters, 2007).

By contrast, the studies of the second category (serious games primarily targeting vocabulary learning in a second language) use the following theories:

While in the previous category there was quite a lot of coherence about the importance of the phonological paradigm, there were much bigger variations within this category. Some of these studies focused to a higher degree on theorization around language, while a few focused more on theoretical considerations about game-design.

On the one end, two studies only relied on language theories and did not rely on game theories at all. Firstly, Rankin et al. (2021) focuses on the concept of morphological awareness as described by Carlisle (Carlisle, 2000; Carlisle & Kearns, 2017), arguing that morphological problem solving skills have been proven to be instrumental in vocabulary growth of second language learners, and considering vocabulary and morphological awareness together as critical for reading skills (Perfeti & Stafura, 2014). They also employed the concept *Willingness to communicate* (Dornyei & Ushioda, 2009; Reinders & Wattana 2012), through which they identified the problem that learning may be hampered if learners are less likely to interact with one another, particularly if a second language learner was expected to interact with a native speaker (Peng, 2014). Rankin et al. (2021) see a potential in game-based learning with video games in that they may potentially be used to increase confidence among learners and encourage them to interact with others. They therefore based their game on this idea. Vasileiadou & Makrina (2017) considered several aspects of vocabulary learning when designing their study: intrinsic properties of words such as pronunciation, morphology etcetera are to be taken into consideration (Takac, 2008), but these are usually overshadowed by challenges associated with extralexical factors such as students' previous language experience, attitudes and motivation (Osguthorpe & Osguthorpe, 2009)

Two studies were theoretically inspired by the RETAIN-model (Gunter et al., 2008), which stands for Relevance Embedding Translation Adaptation Immersion and Naturalization. A central idea with this model is to embed visual elements into a serious game in order to "lure" the player into engaging with for instance a narrative and learn from it. Ensmann (2021) tries to combine this with the concept of visual literacy, considering it of interest to second language learning (e.g. Chen, 2006). Ensmann (2021) also uses the concept Digital Media Literacy (DML), referring to Kenny (2011) who claims that this has to do with being able to discern relevant elements of a plot with regard to time and place, cause and effect and the communication of the plot. For Ensmann (2021), the concept of Digital Media Literacy means that storytelling elements is something that designers within the field of game-based learning need to consider. Ensmann additionally takes the view that

game-based learning can also provide scaffolding, a fail-safe environment (Justice & Ritzhaupt, 2015, Kenny, 2011), and can improve focus and cognitive processing (Eichenbaum et al., 2014; Gee, 2008). The design of Calvo-Ferrer's (2020) game *The Conference Interpreter* was, just like Ensmann's (2021) game, built on the RETAIN-model (Gunter et al., 2008) but, Calvo-Ferrer (2020) also considers the theories that fun-related aspects of game design may distract players by increasing extraneous cognitive load (Paas et al., 2003)

The study of Dore et al. (2019), the only one of this category that focused on first language vocabulary acquisition, based their game design on four principles on the "science of learning" (Hirsch-Pasch et al., 2015): first the ambition that the child should be actively involved with the game with a "minds-on" approach (accomplished in the game by second person narration), secondly, the ambition to create engagement and minimize distractions (accomplished by using attractive images but avoiding extraneous hot spots or sound effects), third that the content should be embedded in a meaningful context in order to optimize learning (accomplished by introducing vocabulary as part of a narrative) and lastly that social interactions will foster better learning (Dore et al., 2019)

### Studies of the third category (serious games built around a narrative) use the following theories:

The two studies of this category used theories from the field of visualization literacy and argument literacy, and combine them with theories on the importance of narration for semantic learning. Hyunh et al. (2021) consider visual literacy as the skill of correctly understanding and interpreting information presented through visual models such as for instance pie charts (Taylor, 2003; Boy et al., 2014; Chevalier et al., 2018) Lawrence & Sherry (2021) employ the concept of *argument literacy* (Graff, 2004) in their study. In accordance with Graff's (2004) ideas, argument literacy is the ability to utilize available information in such a way as to be able to provide convincing and effective argumentation, a skill considered particularly valuable and prioritized in its written form (for instance by American curricula). Factors such as audience awareness and integration of opposing viewpoints is considered particularly important according to research on argument literacy (e.g. Felton et al., 2015, Newell et al., 2011)

Additionally, Hyunh et al., (2021) consider narrative elements of games to have benefits in the form of increased immersion (Schneider, 2004), and heightening intrinsic motivation (Gee, 2003; Liebermann, 2006; Wang & Singhal, 2009). More generally, narration in learning activities is considered to also have effects on memory, alongside increased motivation (Bielenberg & Carpenter-Smith, 1997). Somewhat differently, Lawrence & Sherry (2021) take into consideration the concepts *visual syntax* (the logical sequencing) and *semantics* (meaningful

relationships), referring to Gee's (2015) argument that these are important for understanding how gamers solve problems and accomplish tasks.

### Studies of the fourth category (inductive studies on commercial games) use the following theories:

Since the games of this category were not designed by the researchers themselves, the researchers instead sought to describe or investigate the correspondence between the commercial games and theories that might be applicable. All three studies had a rather inductive approach, where language learning opportunities were sought. For instance, while Vasquez-Calvo (2018) framed their approach using *New Literacy Studies*, specific language strategies or aspects were not identified initially (Vazquez-Calvo, 2018). However, the literacy practices under scrutiny were identified as, among other things, fandom, workflow, and language learning events in the online communities of gamers (Vazquez-Calvo, 2018).

Hobbs et al., (2020) states that the pedagogical idea behind the Minecraft Club is called *the Science Hunters approach* which is built around adults scaffolding as well as collaborative learning approaches (Vygotskyj, 1978; Mercer & Littleton, 2007). Another principle, particularly important for students with cognitive disabilities, is non-competitive learning environments where it is safe to make mistakes (Cassidy, 2013) and where there is a strong feeling of acceptance (Ringland et al., 2016). Hobbs et al. (2020) also fetch inspiration from the field of game-based learning from scientists such as Gee (2007), O'Sullivan et al. (2017) and Prensky (2001).

More deductively, Azman & Dollsaid (2018) took as the point of their departure the 16 learning principles of "good games" as identified by Gee (2003). They sought to identify whether these principles, as suggested by Gee, could actually be validated through their empirical study.

### In summary, the theories most and least common are:

The theory of The Simple View of Reading (Gough & Tunmer, 1986) is the most commonly referenced theory in this review, highlighting the importance of developing decoding skills either through developing phonological awareness, whole word reading, or visual-cognitive training when the children are around the age of school entry. Among game theories, the RETAIN-model (Gunter et al., 2008) influenced two studies, and the ideas of the researcher Gee (such as the learning principles of "good games") are referenced in a number of studies, but otherwise game theories are used much less coherently, with studies referring to many different theoretical elements as relevant.

## 4.2 DESIGN OF VIDEO GAMES USED FOR LANGUAGE LEARNING

The studies of the first category (serious games targeting fundamental literacy skills or pre-reading activities in the first language) use the following designs:

A number of studies in this review sought to evaluate the effectiveness of serious games primarily focusing on the attainment of fundamental literacy skills, focusing particularly on the attainment of the alphabetic principle and the development of phonological awareness. Notably, they all employed experimental design (with differing levels of randomization).

Several studies researched the effectiveness of the Finland-based serious game GraphoGame from a number of perspectives (Baker et al., 2017, Ronimus et al., 2019, Thomson et al., 2020). GraphoGame is a relatively well-researched product (see for instance Lyytinen, 2021), but given the framing of this review only these most recently written articles were included. Another study by Sánchez-Vincitore & Cross (2021) employed a game called Aventura de la Letras which according to the authors share many design principles with GraphoGame. Another game based on similar principles and theories is Escribo play, whose effectiveness is evaluated in a study by Amorim et al. (2022). Samur (2019) performed a study on a game called Kes Sesi, similarly targeting phonological awareness.

These games were similar in design in many respects. GraphoGame, Aventura de la Letras, Escribo play and Kes-Sesi, were largely based on a multi-choice design where the child (working alone) selects items such as graphemes they believe corresponds with auditory input. For example, in Kes Sesi (Samur, 2019) the central game mechanic chosen was for the children to “cut letters” which was performed by the children sliding their finger across the screen on a letter which they believed corresponded with the phoneme heard. The child would then be provided instant feedback (on whether the selection was correct etcetera). In all cases the games also provided gradually increased difficulty. Finally, the games contained thematic elements such as animals in order to provide a playful playing environment (Samur, 2019; Baker et al., 2017; Thomson et al., 2020; Amorim et al., 2022; Sanchez-Vincitore & Cross, 2021; Ronimus et al., 2019).

With GraphoGame (possibly on account of it being a more researched game than others) some additional considerations were made. The articles on GraphoGame point out that that game analyzes the number of correct inputs as well as respond

time. This allows the game to focus specifically on the learning “bottlenecks” of the individual child by targeting areas of difficulty and thus avoiding unnecessary time loss by training skills that has already been automatized. However, in order to avoid discouraging the child, only 20% of the tasks subsequently put forward are ones that are based on content where the child has previously repeatedly made mistakes (see for instance Lyytinen, et al. 2021). Additionally, Ronimus et al. (2019) argues that a game such as GraphoGame is intended to be played in short bursts, over an extended period of time, on the grounds that longer sessions than that could lead to boredom which would possibly discourage the child from playing the game any further (Ronimus et al. 2019). Children developing normally are supposed to use the game for a limited period, but children with learning difficulties may be advised to play the game all the way up to somewhere around the third grade (Lyytinen et al. 2021).

By contrast, built on the theories of cognitive attention, Bertoni et al. (2021) utilized action games for the Nintendo Wii game console, along with other mini-games with similar features. Unlike Bertoni et al. (2021), who used a combination of serious and commercial games in their study, Pasqualotto et al. (2022) developed their own video game-based cognitive intervention with the goal of improving several aspects of executive functions. The idea was to provide mini-games that functioned as gamified versions of standard clinical exercises that feature elements of action video games. The game was under construction for two years and not only game designers, artists, cognitive or clinical experts, but also players aged 8-14 contributed with input. The game, called Skies of Malawak, was designed to be entertaining and consisted of two main mechanics. “The flight”, a fast-paced game focused around distributed attention and shifts between divided and focused attention, and “The village”, built around slow-paced exploratory movement. In addition to these two main games, the product also contains nine mini-games targeting different executive functions. Other design principles utilized were in-game progression with personalized difficulty adaptation, similar to other games of this category.

Serret et al. (2017) employed the serious game SEMA-TIC, tailored specifically towards children with autism. Unlike the abovementioned games of this category, targeting phonological awareness training directly, the aim of SEMA-TIC was instead to focus on children learning to identify words visually, as logographs (identifying each word as a whole picture). SEMA-TIC employed four different types of games built around visual learning: word-drawing associations, sentence-3D animation associations, logical games with words and logical games without verbal instructions. For each game there was a specific list of words to learn. Game design elements included drag-and-drop mechanics, multi-choice questions and visualized feedback and rewards. The game was also designed to be played with a caregiver or family member who could give verbal or physical guidance (Serret et al., 2017). Additionally, the aim was to teach basic grammatical structures of words.

Furthermore, Serret et al. (2017) claim that since many children with ASD have a rapid learning curve the idea was to provide a vast amount of material (with more than 5000 words featured in the game) so that players could go through more or less of the material at their own pace.

A few other games targeting similar skills were built around the idea of providing a wider array of games and exercises, where the game was centered around a hub of sorts where many different types of exercises could be accessed, with differing levels of difficulty: PBS Kids Island, employed in Schmitt et al. (2018), is aimed at 4-6 year old children and is centered around the building of an amusement park. Schmitt et al. (2018) write that the PBS Kids Island game has been influenced by the fundamental literacy skills identified by the U.S. congress, and heavily promoted within U.S. educational policy. These relate to concepts of print, alphabetic knowledge, phonemic awareness, phonics, fluency and comprehension literacy skills, all considered moderate to large predictors of more advanced literacy skills such as writing and general reading comprehension. Additionally, vocabulary development was considered particularly important to include to cater to the needs of children from low-income households (Schmitt et al., 2018). Similar to other games of this category, there was a principle of gradual progression: the children needed to complete four games on one skill level before proceeding with a more difficult level. The games were designed to be short and focused, with corrective feedback and multiple guesses if mistakes were made (Schmitt et al., 2018). IRead (employed in Serra & Gilabert, 2021) came about as a result of the iRead Horizon 2020 Innovation project, and was designed to include an adaptive component and for users of different languages such as English, Spanish, German and Greek. IRead was designed to support 15 different dynamics intended to support various aspects of language development such as phonological and morphological awareness, common sight words, confusing letters, adjectives, pronouns, negation, passives, prepositions, adverbs, determiners, modal verbs, prefixes, suffixes etc. The game mechanics created were, as pointed out by the authors, very similar in design to traditional second language acquisition pedagogy (e.g. multiple choice, matching, odd out one out). Additionally, the program contained hundreds of texts, with prereading activities and other support structures such as a dictionary of 22 000 words. Language items were decided for each language by a panel of language experts and as a general design principle, features known to be easier for learners were supposed to appear earlier in the game. Similar to GraphoGame, students had to succeed with a set percentage of tasks (in this case 70%) of items on a particular language aspect, with a minimum of three exercises carried out for each, before the student could proceed (Serra & Gilabert, 2021).

The games of the second category (serious games primarily targeting vocabulary acquisition in the second language) use the following designs:

The games of this category were serious games designed with a vocabulary learning objective in a second language in mind. It should be noted that by contrast to the others, Dore et al. (2019) was also included in this category since this study also had a focus on vocabulary learning (though in this case of the first language). The other four studies were targeting second language acquisition (Rankin et al., 2021; Calvo-Ferrer, 2020; Ensmann, 2021; Vasileadou & Makrina, 2017). The games are described in more or less detail. For example, in Vasileiadou and Makrina's (2017) study the games used, and how they were designed, are described vaguely, other than that the games were available on online game sites. However, though the study designs differed in several respects, the game designs had several similarities. Most commonly, the games contained images or other representations of objects that the participants were supposed to link to different words using a multi-choice design. For instance, Rankin et al.'s (2021) serious game revolved around providing the players with themes centered around everyday, concrete objects but also abstract ideas. Content themes were generated randomly from a pool, and after being provided pictures relevant to the theme in question, players were supposed to help each other identify relevant vocabulary and concepts. Similarly, in Ensmann's study (2021), different topics, such as letters of the alphabet, animals, colors and other things commonly found in the children's own environment were identified as fitting objects for inclusion in the game design.

Also, several games allowed for several attempts if incorrect inputs were made. In Dore et al.'s (2019) study, when incorrect answers were given, the child was allowed to hear the definition again and be given a new chance. Additionally, words were presented at three separate times: the first time with a definition, the second time with a forced choice task which demanded active involvement, and a third time where the child was supposed to apply the vocabulary to a new situation. In Ensmann's (2021) game, after three incorrect answers, the child was provided with the correct answer and allowed to proceed.

Calvo-Ferrer (2020) employed a serious game called *The Conference Interpreter*, specifically designed for second language vocabulary acquisition. Though the specific designs are described quite hazily, the game is described as a drill-and-practice game, based on trial-and-error mechanics and instant feedback. The authors concede that these mechanics are a salient feature of the design, more so than in other types of games and that it may not be the most motivating type of game for this reason (Calvo-Ferrer, 2020).

In some cases, the designers tried to incorporate elements of narrative around objects in order to strengthen learning. For instance, in Dore et al.'s (2019) study the researchers wrote the game scripts and the logic of gameplay, while a hired educational development company provided static cartoon-style images for the game. In Ensmann's (2021) study, pre-service teachers wrote digital stories in collaboration with IDT-students who converted the stories to games under the direction of Ensmann.

Rankin et al.'s (2021) study was different from the other studies in a few respects. Firstly, there was a significant communicative element. One of the students was supposed to describe the relevant object to their partner. Based on the theory of WTC (Willingness to communicate) the idea was to offer a lowered fear of mistakes and a heightened sense of anonymity by not having the students physically meet. This was expected to be particularly beneficial given the students' differing language proficiency. Instead, students only communicated via an in-game-chat. Another difference with other game designs was that answers were generated in the form of written chat messages (instead of simply clicking on choices), with the intention of providing both an opportunity for the students to support each other's language development, but also an opportunity for the researchers (and potential teachers) to observe the students' vocabulary development and morphological awareness development. The idea was also that writing would favor participants' reflective thinking and thus their morphological awareness (Rankin et al. 2021).

The idea of differentiation was important for Vasileiadou & Makrina (2017). Their study was structured in such a way that the first game played was chosen by the teacher, after which the students would continue by selecting a game that they deemed relevant. This allowed for a degree of differentiation since the students could choose either a more difficult or an easier game compared to the one initially chosen by the teacher.

**The games of the third category (games built around a narrative) use the following designs:**

Unlike the first two categories, the games of this category were built around the idea of reading and interpreting sentence-level sequences of text, along with visual information such as graphs, rather than focusing on individual letters, syllables or isolated words. This would require a degree of reading fluency, along with an element of interpretation.

The designs of these two games also allowed for the player to control an in-game movable avatar (of the character), a feature largely absent from games in category one and two. However, there was also design similarly to those categories in that

regularly, the players would make multi-choice decisions to provide answers for questions. These would be programmed as either right or wrong.

Huynh et al. (2021) wished to investigate how a serious game with narrative elements, and enhanced with visual elements, would affect content learning and engagement, arguing that games and gamification is an area that has not yet been explored within the emerging field of visual literacy. Huynh et al. (2021) employed a so-called narrative-focused role-playing game. Two versions of the game were evaluated, one with narrative elements and one without, with the purpose of investigating whether the inclusion of narrative events affected semantic learning gains. In accordance with aforementioned theories, the game was based on principles of visualization and education science but also general game design principles and narratives in games. The game was designed around a linear narrative, in other words a simple story with a dramatic arc. There was also an ambition to tie the encountered problems, in the forms of questions presented to the player, to the story. The purpose of the narrative was also to be engaging enough that the students would be inclined to keep playing even when more difficult problems were encountered towards the end of the game. The game interface shifted between three modes: activity view (with chart-related questions), a dialogue view (where the children would read game dialogue) and an exploration view (where the children could control their avatar) (Huynh et al. 2021).

Lawrence & Sherry (2021) argue that compared to the relationship between videogames and reading, relatively little research has been done on the relationship between videogames and writing. Additionally, they argue that opportunities are plentiful to investigate the pedagogical quality of feedback provided by videogames, which in turn inform students' thinking and arguments. For this reason, they decided to evaluate the free online educational game *Quandary*, designed to teach argumentation to players ages 8-14. The idea behind the game was to teach the students about different arguments regarding environmental protection which would ultimately materialize in their own communication through writing. Input and feedback from the game (Lawrence & Sherry, 2021) included both oral and written discourse but the researchers noted that it would also offer visual and auditory cues, for example body language and intonation of in-game characters. The idea was that players should draw reactions from the characters of the game via spoken as well as written input, with the students' own input in turn eliciting feedback from the game. The characters were intended to represent different vocations and consequently different perspectives and viewpoints on environmental problems (Lawrence & Sherry, 2021). With regard to game mechanics, the game revolves around the player categorizing characters' statements under headings such as "fact" or "solution". These categories have been coded as either right or wrong, and the player must make a set number of classifications as "correct" before the game allows the player to continue (Lawrence & Sherry, 2021).

The games of the fourth category (commercial games) use the following designs:

The games of this category were different than the other categories in that they were so called commercial games, which meant that their designs could not be considered in their entirety in these studies. Instead, the studies had to identify relevant design elements of interest from a language learning perspective. Similar to category three, the games featured avatars (controllable characters), sentence-level sequences of text, and multi-choice designs. A notable difference from the other categories was that these game designs allowed for interaction between different players.

The content examined by Vasquez-Calvo (2018) stems from the language content of games that the participant chose to interact with (for example the narrative being presented in text boxes), but which is then interpreted and handled by the participant of the study. A game which is presented as an example is called *Master of Darkness*, which supposedly makes use of literary cues from the likes of Dracula and Jack the Ripper and generally fetched inspiration from mystery murder stories. The activities analyzed was translation work but also the language discussions on internet forums (Vasquez-Calvo, 2018).

Hobbs et al. (2020) studied how the commercial version of the game Minecraft was used in a “Minecraft Club”. According to them the semi-formal setting meant that learning objectives were relatively loosely defined compared to what would have been the case in a formal school setting. This materialized in that learning objectives were still centered around a particular science topic, which was intended to provide a clear focus for the session and signpost the imaginative building and exploration that the children engaged in. Examples of such topics were renewable energy, plant propagation, and low-temperature environments. The children were also encouraged to provide suggestions for topics that they would like covered during the sessions. Another highlighted purpose of the club was to foster communication. For this reason, it was considered vital that the children and their guardians were gathered in the same physical locale despite playing a digital game. This was a deliberate choice since both parents and therapists (the children who participated in the club were mostly diagnosed with ASD, ADHD and/or dyslexia) wished to enforce connections between virtual and “face-to-face” (or at least being in the same physical place) communication and relationship building (Hobbs et al., 2020).

Azman & Dollsai (2018) utilized the popular game World of Warcraft on the grounds that it was the leading subscribed MMORPG at the time of the study. They collected data from questionnaires, in-game chat logs and email interviews from five participants playing the game. A content-analysis technique was used on qualitative data in the form of data from questionnaires, in-game chat logs and email

interviews. Azman & Dollsaid (2018) wished to explore how the literacy practices of online gamers influenced their language learning through Massively Multiplayer Online Games (MMORPGs), which they regard as a type of serious game with learning potentials for EFL (English as a foreign language). They argue for their study on the grounds that MMORPGs may have pedagogical value, but a greater understanding of the design of these games and potential uses needs to be explored. More specifically, the authors wished to investigate the communication behaviors of speakers as they materialized in the game.

**In summary, the most and least common game designs are:**

The strongest tendency within game designs of all studies, particularly of the three first categories which include serious games, is to base gameplay on a multi-choice design, in other words the participants selecting from a number of potential answers and then choosing one. If this answer is correct, the player is allowed to proceed. The purpose of this is to connect different language elements (such as phonemes, syllables, whole words) with a visual representation such as a letter or a word. This design is prevalent in category three and four as well, but in that case the players must interpret sentence-level sequences of text and present a similar reply. Only in a very few studies were the players supposed to write answers. Avatars, allowing for a degree of exploration, was found predominantly in categories three and four. Designs providing interaction between players was only identified in four studies, of which three was found in the fourth category.

### **4.3 LEARNING GAINS**

The studies of the first category (serious games targeting fundamental literacy skills or pre-reading activities in the first language) reported the following learning gains:

The studies of the first category of games were all experimental studies, which means that they drew conclusions about the effectiveness of the included games. Several of these games report positive results particularly with regard to different aspects of fundamental literacy development, such as phonological awareness and letter recognition. Note that though these studies measured slightly different subsets of skills, they are all related to the phonological paradigm. For example, Ronimus et al. (2019) measured word reading which was expected to improve as a result of better decoding skills.

Some studies report very positive effects on reading abilities such as phonological awareness, decoding abilities and letter knowledge. Amorim et al.'s study (2022), evaluating the effect of the game *Escribo play*, showed particularly strong significant improvements on both the reading and writing abilities of the children in their experimental group compared to their control group. The experimental group developed their reading ability over three times more than the control group, and over two times more than the control group with regard to writing. The authors also collected qualitative data from the teachers, who similarly reported that from what they could see, many of the students had evolved a lot. Amorim et al. (2022) are thus confident that their researched game, *Escribo play*, provides a very effective way to practice fundamental literacy skills such as phonological awareness.

Schmitt et al. (2018) found that children in their experimental group performed better than those in the control group on a number of measured literacy skills, among them being able to name more lowercase letters, knowing more letter sounds and identifying letter sounds more rapidly. Among the significant results, effects were particularly strong for phonological awareness. There were also moderate gains regarding vocabulary. They are particularly confident in their results given that their study utilized a randomized experimental control group design, claiming that much of the controversy around these kinds of studies previously has been their lack of rigorous methods (Schmitt et al., 2018). Serra & Gilabert (2021) also reported positive results. As a matter of fact, they also noted that whether the sequence of game activities was designed by an experienced teacher or the game program itself did not significantly affect how much the players improved. Both groups improved their overall accuracy and fluency of reading. This despite significant differences also in how players approached the program depending on which group they were in. However, playing a narrower variety of features did lead to greater gains in decoding skills. Another finding was that a weak-to-moderate but significant correlation was found between level of initial word decoding and gains in decoding. In other words, poorer decoders at the start of the study improved the most. This relationship could also be seen when testing decoding skills through nonword reading with a moderate and significant improvement, and in this case students of the algorithmic group improved more. The same was true when testing reading fluency, but again only in the case of the algorithmic group: the less words per minute learners could read at the beginning of the study, the more fluent they became. Additionally, learners who displayed an initial worse listening ability also made more gains during the project. Finally, the results of Samur's (2019) study showed that there was a small but statistically significant difference in favor of the experimental group with regard to sound recognition. The results of Thomson et al.'s (2020) study showed that neither second language status, nor gender, explained any significant variation in growth when playing *GraphoGame*, which means that the game can be used equally successfully by participants regardless of these factors.

When evaluating the game SEMA-TIC, Serret et al. (2017) found certain transfer effects since improvements were also seen in novel materials such as alphabet knowledge, literacy conventions, and decoding (since, unlike other studies in this category decoding was not actually targeted during the intervention) In the end, the authors conclude that three out of twelve children in the experimental group could be considered word decoders by end of the intervention, compared to none in the control group. The researchers take the results to mean that game-based activities centered around whole-word reading can also be a viable alternative to the phonological paradigm, at least when training children with autism (Serret et al., 2017).

Decoding abilities were also successfully improved through games built on theories of cognitive attention. Bertoni et al.'s (2021) results indicated that after their intervention attentional control and phonological decoding improved. On the other hand, they also noted word reading performance did not improve, meaning that similarly to Ronimus et al. (2019) they interpret their results somewhat cautiously. However, Bertoni et al. (2021) conclude that their study supports the theory that design features of certain action games, such as the speed of moving objects which demands a high degree of perceptual and motor load, as well as peripheral processing, stimulate the same parts of the brain that are used when performing reading tasks such as reading pseudowords (Bertoni et al., 2021). This in turn is a skill that is linked to phonological awareness and decoding skills. Based on similar theories and designs, the game study of Pasqualotto et al. (2022) showed that both reading abilities, as well as attentional and planning skills were significantly improved during the study. Playing the game also reduced reading errors. In addition, they noted that younger children improved more. During a follow up-test six months after the study, the improved reading efficiency had been maintained. Interestingly, the authors also noted longer-term effects: when comparing the grades in Italian (the participants' mother tongue) for the experimental group with the control group, there was a small advantage for the former group (Pasqualotto et al., 2022).

By contrast, a few other studies of this category report that while they see specific improvement with regard to particular skills, they are careful not to draw conclusions about reading generally since expected transfer effects did not always materialize. However, all these researchers also point out that there may be particular explanations for this. The results of Ronimus et al.'s (2019) study indicated that their six-week intervention with GraphoGame did indeed give a boost to the word reading of second-grade students with reading difficulties. However, Ronimus et al. (2019) identified no immediate transfer effects on spelling, sentence level reading fluency or reading comprehension. Therefore, Ronimus et al. (2019) interpret the effectiveness of GraphoGame somewhat cautiously, though this is with regard to children with cognitive disabilities, which may struggle more than others. That children with cognitive risk status at school entry made less progress than other children was also shown in the results of Thomson et al.'s (2020) intervention.

Sánchez-Vincitore & Cross' (2021) study showed that the experimental group outperformed the control group with regard to syllable recognition after the intervention was carried out. However, similarly to Ronimus et al.'s (2019) study, transfer effects to other aspects of reading abilities were not identified, meaning that results must be interpreted cautiously. On the other hand, the authors point out that the short duration of the intervention may also explain this lack of transfer. Baker et al.'s (2017) study on GraphoGame did not show that students in the experimental group had developed significantly better decoding skills than those in the control group. However, the researchers theorize that it is possible that the teachers of the control group may have spent particular attention to training phonological awareness during the intervention.

**The studies of the second category (serious games primarily targeting vocabulary acquisition in a second language) reported the following learning gains:**

While a few studies in this category reported positive effects on vocabulary learning in a second language, evaluative measures differed to the degree that it is difficult to draw general conclusions about vocabulary learning gains through the use of video games.

Two experimental studies reported positive results. Dore et al. (2019) demonstrated that children in the experimental group showed significant gains for target words. This result was valid for both a receptive as well as for an expressive test, indicating that the children had also learnt to use the words in their own production (Dore et al., 2019). Vasileiadou & Makrina's (2017) most significant result was that, after 24 weeks, the students who had achieved the lowest scores on vocabulary in the pretest were the ones who improved the most in the post-test. By contrast, the improvement (as measured through the post-test) of those students who achieved higher results on the pre-test was limited.

By contrast, the results of Rankin et al. (2021) showed that the outcomes of the different student pairings differed significantly with some students interacting significantly more than others (and in some cases some students leaving interactions altogether) meaning that some students benefited to a much higher degree from the experiment than others. The authors conclude that vocabulary learning was limited (Rankin et al., 2021).

While Calvo-Ferrer (2020) employed an experimental design, his comparison was rather between students dependent on their technological affinity using the digital nativeness scale. Calvo-Ferrer (2020) could not see that technologically savvy students will benefit more from game-based vocabulary learning than others (Calvo-Ferrer, 2020) since there weren't any major differences between students.

However, the factors "comfortable with technology" and "thrive on instant gratifications and rewards" were statistically significant predictors, which lead the author to add that there may be small differences between learners. Inversely the factor "able to multitask" had a negative impact on learning outcomes, indicating that students who are easily distracted may not learn as much from video games intended to promote the learning of vocabulary (Calvo-Ferrer, 2020).

Finally, Ensmann (2021) did not quantitatively evaluate learning gains, but evaluated her study using Weston's formative evaluation model (1995) to gauge the effectiveness, efficiency and appeal of the instructional materials in order to inform design of games focused on young children's second language learning. Also, a so-called facial expression reader (Noldus et al., 2001) was used to gauge reactions from testers. These types of data were complemented by testers also answering open-ended questions. Additionally, the students involved in the project carried out observations of testers playing the games. Conclusions drawn from evaluations were the following: first, it was deemed important that characters/avatars in the game needed to be relatable for the children to be engaged (for example, a game centered around the concept of "gems" was deemed unsuccessful since the concept of gems seemed too foreign to the children). Secondly, the games should offer cues, since there were instances of the players becoming lost without them. Thirdly, navigational elements needed to be clear and offer choice. Fourth, the student designers recognized the importance of imagery, sounds and "winning" stimuli in order to evoke reactions from players (Ensmann, 2021).

### **The studies of the third category (serious games built around a narrative) reported the following learning gains:**

The results of both studies were mixed to negative. While both studies were designed as experiments, similarly to the vast majority of the studies of first two categories, the qualitative dimension was important in this category since the two studies interpreted many different situations arising throughout gameplay. It was also through qualitative means that the researchers took note of a number of problems arising during and after gameplay.

Hyunh et al. (2021) drew the conclusion that while the inclusion of narrative elements in a game doubled the amount of time needed to complete the game, there were no significant differences regarding semantic learning between the experimental group and the control group. Lawrence & Shelly (2021) similarly reported that the participating children seemed to have misunderstood a lot of the content in the game, given how they ultimately argued in texts after playing the game Quandary. When taking part of the children's written argumentation after playing the game, the children ultimately do not recommend the same types of solutions that they were provided in the game.

While playing *Quandary*, it became clear that observed children did not fully understand the language (or semantics) of the game. They resorted to strategies such as listening for cues, or outright guessing, in order to advance. Particularly when students encounter situations in which there is only a single correct answer, the researchers noticed they may try to circumvent the task (Lawrence & Sherry, 2021). Hyunh et al. (2021) made similar conclusions about this aspect. Interview answers in their study generally indicated that there seems to have been problems with matching the difficulty level of the game to the participants' pre-knowledge, with many participants reporting that pie-charts and visual aids might be more helpful to someone who was younger than them or to someone for which the content was very unfamiliar.

Lawrence & Sherry (2021) also conclude that part of the problems may arise from design decisions. Confusion may stem from the game presenting different solutions as equal, leaving the students to interpret the arguments themselves and drawing very different conclusions depending on which game ending they received. Additionally, Lawrence & Sherry (2021) point out that since the feedback from the game rewards or penalizes predetermined "right" or "wrong" choices, this may lead to students developing argumentation strategies that are relatively ineffective, while at the same time fostering a false sense of competence among the students (Lawrence & Sherry, 2021).

On the other hand, a potential strength of narratively driven games, identified by Hyunh et al. (2021), is that they may heighten engagement levels among participants. They noted that the level of engagement from participants was significantly higher when narrative elements were part of the activity. Participants of both groups also expressed desire for more story and interactions between characters (even the game variant with a narrative was estimated to take only about 30 minutes to play through). Another possible advantage of the narratively driven design is that while observers noticed a number of cases where participants seemingly made selections without properly considering alternatives, making use of accompanying instructions, information etcetera, this was mostly visible in the group without narrative elements. On the other hand, even in the group with narration, there were a few participants who displayed signs of fatigue (despite the game only being around 30 minutes). Another result was that when evaluating the number of correct answers, male participants performed worse than in the control group. There were also conflicting and highly varied reactions on the game feedback.

**The studies of the fourth category (commercial games) reported the following learning gains:**

Notably, the games of this category seemed to a higher degree than the others stimulate pragmatic, or social, aspects of gameplay. Participants of all three studies

consistently played and interacted with others either through speaking or writing. The researchers point out that this allowed them not only to develop pragmatic skills but it also provided a supportive environment in which the players were able to provide and receive feedback on their language use, thus offering opportunities for improvement in other areas such as vocabulary and morphological awareness. However, it should be noted that since these studies were designed as case studies, these results must be cautiously analyzed.

Azman & Dollsaid (2018) identified a number of situations in which participants' language skills may have improved as a result of the analyzed literacy practices. Firstly, they discovered that their participants were able to develop their communicative competence through exchanges and collaboration. They argue that the observed players feel safe to continuously make queries of concepts and words appearing in the chat. The effort to reach collective goals also allows for player bonding which the authors regard as an important motivating factor. Secondly, content matter such as words and concepts relevant to the game (and the outside world) is experienced and possibly internalized. The situations provided by the game are regarded as authentic and contextualized, and the player is forced to be linguistically active in order to proceed in the game. Situations demanding clarification and comprehension checks arise which allow for learning opportunities. With support from other player characters, but also non-playable characters (that is, computerized characters found in the game), Azman & Dollsaid (2018) suggest that the players are able to solve problems they otherwise wouldn't be able to. Thirdly, they can see that players are able to provide corrective feedback to each other, for example regarding bad typing which also provides opportunities for players to improve finer details of their language skills such as grammatical competence and spelling. A reservation is made, however, in a concern over some of the language used by players. For instance, there is frequent usage of abbreviations, and spelling mistakes are common. On the other hand, these imperfections allow for opportunities for improvement and use of communicative strategies. Azman & Dollsaid (2018) argues that collaboration and cooperation between participants generally means that they can overcome these types of issues in order to make the intended meaning clear. The authors conclude that through their literacy practices while playing an MMORPG like World of Warcraft players develop their pragmatical language skills such as communication and negotiation skills, but also vocabulary, listening skills and to some degree reading skills (Azman & Dollsaid, 2018).

Vasquez-Calvo (2018) concludes that his participant has managed to learn a complex set of literacy skills, which can be broadly divided into two categories: language skills and IT skills. Vasquez-Calvo identifies a workflow of seven stages in which the participant uses these skills interchangeably, depending on the stage he is currently in. These literacy skills go hand in hand with the participant's language development. Initial attempts at understanding, interpretation and translating are

later revised through feedback from other people online. This allows the participant to learn meaningful, situated language items. This also has the function of offsetting the participant's own language deficiencies, improving things such as spelling, grammar but also adjusting language use based on considerations surrounding audience, cultural considerations and pragmatically adequate solutions. Vasquez-Calvo (2018) notes that most of the feedback concerns typography, spelling and grammatical correctness, but there are also instances of feedback regarding lexical-syntactical issues and to some degree feedback regarding socio-pragmatic and socio-cultural aspects. The participant gets feedback on his language both regarding choice of words, unclarities, register and tone, and efficiency. The participant must consider linguistic form, semantic meaning and pragmatic use not only of his own language but also other languages that he is translating from (Vazquez-Calvo, 2018).

In Hobbs et al.'s (2020) study, the participating children and parents were asked to evaluate whether the children's participation in the Minecraft Club had resulted in some form of learning, and if so, to more specifically identify what that would be. 88.5 percent of the children participants answered favorably to this question, and while there were some that pointed out the science content, the most common answer was that participation in the Club had fostered social and communicative skills, in other words pragmatical aspects of language. In interviews, parents also attested to their children becoming more communicatively active and more inclined to take social initiatives and engage in communication with other children. Though such self-reported learning must be interpreted very cautiously, the results come across as strikingly positive. The authors draw the conclusion that the Minecraft Club-setting provides a context in which children can develop their communication skills, their social skills, build confidence, while at the same time learning (to a degree) about science in a semi-formal educational setting (Hobbs et al., 2020).

**In summary, the most positive and negative learning outcomes are:**

Among the most positive results, studies of the first category strongly suggest that video game-based learning can be used to improve the pre-reading skills (also called fundamental literacy skills in this study) of children about to learn how to read (around the age of school entry). A few studies have some reservations about the question of transfer with regard to these types of games, however, since improvements could only be demonstrated in very specific skills in those studies. Case studies on commercial games also indicate positive results, demonstrating a great potential of the video game medium, particularly with regard to the pragmatic or communicate aspect. A number of studies are able to demonstrate positive improvements on vocabulary of both first and second language learners, though

studies had very different methodological and theoretical approaches. By contrast to these positive outcomes, serious games built on mainly narrative designs reported mixed to negative results.

# 5 Discussion

## 5.1 Comparisons with previous research

This review differentiates itself from the reviews listed in the background chapter with its view that language learning consists of distinct dimensions such as phonology, morphology, pragmatics. This *definition* of the subject matter (Wolfswinkel et al., 2013) was translated into a search strategy which informed the outcome of the study. Consequently, this definition may have highlighted research mostly focusing on young children's first language learning. By contrast, previous reviews are situated mainly in an L2 language learning or an ESL (English as a second language) framework. It seems that when focusing on vocabulary, studies tend to involve older participants such as high school and university students, and involve other types of games such as MMORPGs.

For example, in Hung et al.'s (2018) review it was noted that studies were characterized by a focus on the language learning of university students and most studies focused on MMOPGS. In this review, there was only one such study (Azman & Dollsaid, 2018) and generally the vast majority of studies involved serious games rather than commercial games. A reflection around this is that it is important to explore different areas of language learning since these correspond to varying degrees with the needs of different groups.

One area in which the results of this review strongly correspond with prior reviews has to do with research design. Rosman et al. (2013) and Hung et al. (2018) noticed that methodologically, there is a strong tendency to employ mixed methods within research in this field. In this review, although the vast majority of studies were designed as experiments (with different levels of randomization) almost all of them incorporated qualitative dimensions. A reflection that can be made from this is that while quantitative outcomes are generally desired in order to provide evidence, it is not enough: people's perspectives and interpretations, as well as situational aspects are of great interest within all studies and must be taken into consideration when considering how game-based learning may be utilized in different ways.

Hung et al. (2018) & Xu et al. (2020) concluded that there was a heavy emphasis on vocabulary learning in their reviews. While vocabulary was the most relevant language aspect in a number studies in this review as well, several studies handled

the topic more peripherally. For instance, Rankin et al. (2021) focused on how increased morphological awareness could benefit vocabulary learning, and Calvo-Ferrer (2020) explored how digital nativeness may be indicative of language acquisition. Thus, the results of this review indicate that there seems to be a rather large amount of uncertainty about how to approach vocabulary learning through video games, both theoretically and with regard to design. Though some experimental studies report positive vocabulary learning gains, more research is needed to find stronger connections between approaches and results.

Instead, this review greatly emphasizes studies on games training fundamental literacy skills such as phonological awareness, with 11 out of 21 studies placed in this category. The great significance of phonological awareness is strongly emphasized in research on children's language development (see for instance Lundberg, 2010), since it is considered a fundamental skill for the eventual development of general reading comprehension skills. However, this dimension is largely absent in previous reviews on the connections between game-based learning and language learning. The results of this review also indicate that improving children's phonological awareness may be one of the most useful ways to employ game-based learning for language learning purposes. They also suggest that there is merit to the statement that not only theories highlighting a phonological approach to learning, such as the Simple View of Reading (Gough & Tunmer, 1986) but also theories from the field of attention control may be successfully combined with game-based learning in order to promote phonological awareness.

A common argument for game-based language learning (often referred to in included articles) is that the differentiated, or individualized content is better adapted to students with different knowledge levels. However, the results of this review indicate that currently, the types of games that are able to capitalize on this potential strength seems to be games targeting fundamental literacy skills and commercial games. Several researchers involved in the first category of studies (serious games targeting fundamental literacy skills) report that the ability of these serious games to employ designs that adapt to an individual child's cognitive level, need for repetition, individualized work tempo etcetera is probably an important factor in explaining the identified learning gains (see for instance Baker, 2017). In their review, Xu et al. (2020) criticized games studies generally for not reporting the language level of participants, and advised a higher use of commercial games since they found that custom-made games (or serious games) struggle to match the language level of the users. In this review, the serious games trying to employ a coherent narrative did indeed suffer from these kinds of problems, as reported by the researchers. A reflection that may be made from this is that at least for now, when trying to employ stories and narratives to foster language learning, commercial games may possibly be a better alternative than serious games, in accordance with Xu et al. (2020) ideas.

Another potential strength of game-based learning, as indicated by the results of Rosman et al. (2013) review, is games' potential in heightening motivation. The motivational aspect is certainly prevalent in many of the studies of this review as well. However, the results indicate that this point must also be nuanced, at least when it comes to serious games. A number of articles of this review report a very positive initial response from participants when introducing serious games (and thus not only commercial games are seen as initially appealing to children). However, a number of researchers of serious games report a decline in interest after a number of weeks had passed. The big question then is if this "honeymoon period" with a game is enough to warrant meaningful learning that leads to relevant transfer effects and long-term learning. The answer is probably that it depends on the learning goals of the game: for example, a ten-week game-based intervention (if we assume the child may become too unmotivated to continue after ten weeks) may be enough to improve phonological awareness, but not a sustainable method to improve vocabulary. As illustrated by the results of Ronimus et al.'s study (2019) it is also important to note different aspects of motivation, in the form of emotional or cognitive engagement, with the latter being what is crucial for learning to happen. As pointed out by Ronimus et al (2019) it should also be noted that increased motivation may not necessarily lead to better learning outcomes (for example Annetta et al 2009; Kim et al 2017; Ronimus et al 2014; Wrzesien & Raya 2010).

## 5.2 Theoretical contributions

When we talk about game-based learning, what do we actually mean by the word "game"? On the one end of the line, we have the kind of enormously popular, mechanically complex commercial games (like Minecraft or World of Warcraft), designed mainly for entertainment, that people spend a significant amount of their spare time playing (games from category four of the review). At the other end of the line, we have serious games, designed by researchers, based on highly focused, even simplistic designs (games from category one but to a degree also category two and three) Do all these "games" really have anything in common? This is a question that could warrant further inspection, given that in the articles included in this review, the concept of "game" is seldom problematized or discussed.

It can also be argued that the concept of game-based learning is problematic because there is no absolute distinction between games and other types of digital interventions or programs. Particularly, it can be discussed if the games from the first category identified in the review should be labelled as games. The researchers themselves already use other terms in tandem with the "game"-term. For example, GraphoGame, a game researched through several articles in this review, is described by Thomson et al. (2020) not only as a game but also a *digital reading intervention*. Amorim et al. (2022) describe Escribo play as a *game-enhanced*

*educational phonological and phonemic awareness program*. Another term used by Baker et al. (2017) is *CARI: Computer Adaptive Reading Intervention*. The point of such names is of course to highlight relevant aspects of the design. Somewhat problematically for the concept of game-based learning, these studies were the ones that could display the most positive results in the review. It should also be noted that those types of games don't take much inspiration from the theories of game design but rather that of literacy, meaning that it is difficult to connect theories on game-based learning to positive learning gains. Consequently, the question is how much the "gaming" elements actually contribute to learning, or at least have some other function to play such as increasing a child's motivation to engage in the learning activity.

A related problem when considering game-based learning as a concept is the relevance of different game-design elements. It is important to note that even those researchers which report favorable results in this review point out that there is always a degree of uncertainty about what elements actually contribute to learning (see for instance Amorim et al., 2022). If we consider GraphoGame as an effective learning tool for learning decoding skills (Lyytinen, 2021), does the fact that we might consider it a game have anything to do with its effectiveness? If so, what are the critical features? It seems reasonable to assume that different aspects of design are relevant depending on what the game is trying to accomplish (as related to different theories). For example, Bertoni et al. (2021) identifies the intrinsic characteristics of action games, such as the speed of moving objects, as relevant when targeting decoding skills among children with dyslexia, in line with the theories on attention control. When considering the design of commercial games, Azman & Dollsaid (2018) consider the 3D graphical environments of a game like World of Warcraft, with its role-play features to be of relevance to enable contextualized interaction and authentic language development, in accordance with Gee's (2003) theories. For Ronimus et al. (2019) it is the repetition and drill needed to master certain skills that is the main advantage of a game like GraphoGame, since this increases the exposure time spent on completing learning tasks. Repetition is also an important feature of some of the vocabulary learning games, such as in Dore's (2019) study. Schmitt et al. (2018) reported that in their study, the element of repetition seems to have mattered more for more fundamental literacy skills such as phonics and uppercase letters, compared to more advanced literacy skills. Amorim et al. (2022) point out that for a game like *Escribo* play, very specific considerations regarding activities, dosage, sequences, frequency and the most effective instructional strategies were considered relevant, and that this was only possible through quantitative and qualitative findings from several large-scale meta-analyses.

This study endeavored to not just investigate the effectiveness of game-based learning on language learning on a general level, but attempted to connect learning gains

more closely to theories and design elements used. It is my view that future research efforts should strive to continue this endeavor.

### 5.3 Practical implications

One of the most important outcomes of this review is the realization that well-designed games can help children develop critical, fundamental literacy skills in situations where pedagogical and parental guidance may be lacking. Schmitt (2018) argues that there are strong incentives to provide free interventions for children with generally less access to these types of interventions. Amorim et al. (2022) stresses that there is a great need to develop these types of interventions for children in poverty. A very interesting result of their study (2022) is that the considerable effect sizes of their intervention happened *despite* rather severe, unforeseen problems arising during the study. These were staff shortages (staff expected to aid the children when needed were not available as planned), natural disasters and teacher strikes (preventing the children from coming to school and participating in the study for as many days as was originally planned), and ICT deficiencies (problems with maintaining online connections). Schmitt et al. (2018) also claim that their study provides evidence that web-games designed specifically to teach literacy skills (and based on the same principles) can be effective, even when they are played at home *and with limited support* from teachers, parents or other guardians. Samur (2019) also suggests that their game can be used for self-supported learning and circumstances with very little adult input. Sanchez-Vincitore & Cross (2021) also bring up this perspective, pointing out that in countries such as the Dominican Republic, where their study was carried out, students do not always receive systematic phonics instruction so games like the one employed in their study can still make a valuable contribution. These researchers thus point to a particularly important use for game-based learning: that it may be of particular relevance for developing societies or disadvantaged schools.

Another practical aspect to take into consideration is a teacher who is planning to use video game learning to provide semantic learning through narration and storytelling must take great care since the studies of this review illustrate a number of problems associated with this approach. In accordance with Xu et al.'s (2020) point, it may be worthwhile to investigate how commercial games can be used as an alternative. On the other hand, the results of this review show that while researchers on commercial games point out strong positive potential for these games, they cannot provide much advise on how to employ these games in formal educational settings. Considering the results of this study, the question of how commercial games can be successfully utilized in a formal setting is still somewhat left in the open, since the question is not explicitly explored in any of the articles. Hobbs et al.'s (2020) study could possibly serve as an inspirational example of how a

commercial game like Minecraft can be utilized in a semi-formal setting with high perceived value for participants and parents.

## 5.4 Suggestions for future research

As has been pointed out, the results of this review indicate that more research is needed on how to successfully foster vocabulary learning through video game-based learning. That there seems to be relatively little agreement about how to theoretically or methodologically approach vocabulary learning through video gameplay is in itself something that might be interesting to explore further in future studies. What theories and game-design elements are to be considered most valuable when targeting vocabulary? Utilizing game-based learning built around narration and story is also a potentially promising approach, but only two studies in the review investigated this dimension and with limited success. Therefore, an interesting question is how serious games with narrative elements aimed at, for instance, relaying a particular message, or acting as a sort of immersive reading practice, can be designed more successfully. Alternatively, can commercial games be utilized as for this purpose even in formal school settings, like Xu et al. (2020) suggests?

However, perhaps the most tangible reflection is how highly underutilized the social, or pragmatic, dimension is in included studies. In their review, Hung et al. (2018) pointed out the need for studies to focus to a higher degree on communicative- collaborative- and creative skills. The outcome of this review point in a similar direction. Almost all the included studies attempted to investigate the effects of playing a game in physical or social isolation. The participants were usually sitting in a separate room with only a supervisor present, and with very limited interaction with other players. Alternatively, if others were present, the children would use headphones so as not to be disturbed. For some reason, current serious game design seems highly focused on the games themselves providing all assistance to the player, rather than in tandem with other forms of support from other players or even adults. This also means that whether an intervention is successful or not is almost entirely up to the game design, since little to no support can be expected if problems arise. But would not game design benefit from being more interactive? While the pragmatic dimension, including such aspects as turn taking, social protocol and so on has its own value (particularly to disadvantaged children), interaction between players is also necessary in order to be able to foster collaborative learning or allow for players to help each other solve problems. Such designs could *make use of* interaction as a way to learn or practice other aspects of language such as vocabulary.

In this review the few articles that focused on these types of skills were the ones who, firstly, employed commercial rather than serious games in their study, and secondly, who analyzed language learning among relatively older students. All

three studies evaluating language learning with commercial games focused to a high degree on the social dimension of these games, with Hobbs et al. (2020) particularly highlighting the social and pragmatic learning that happens within the Minecraft Club, and Vasquez-Calvo (2018) and Azman & Dollsaid (2018) pointing out the corrective feedback that participants playing games like World of Warcraft give to each other continuously, promoting improved language use. Unfortunately, since commercial games are largely used outside formal education, there is a significant risk that the social dimension is largely lost when employing game-based learning in school settings. Only a single study on a serious game employed mechanics that were centered around communicative and collaborative elements (Rankin et al., 2021). On the other hand, this study can also serve as an example of the potential problems that can arise when trying to design a game centered around social dynamics. Therefore, future studies should endeavor to investigate how to incorporate, or take advantage of, social dynamics more successfully in serious games. This is particularly true for game designs targeting comparatively young children. Alternatively, studies could investigate how commercial games may be successfully utilized in more formal settings in order to provide students with both training related to the pragmatic and social aspects of language use. As studies on commercial games illustrate, this may also strengthen other language aspects such as morphology or vocabulary.

Other aspects that are of interest are how feedback systems and principles of individualization and adaptation can be further refined to increase learning gains. As illustrated by the results, not all serious game design manages to integrate these principles successfully and adapt to the learner. The risk of this is that the learner will not be able to properly comprehend the learning content, lose interest or resort to brute-force behavior where the player is no longer actually considering the content but just trying to get through the game mechanics. Furthermore, as Ronimus et al. (2019) point out, even well-researched games such as GraphoGame have not reached their maximum potential when it comes to feedback design. Considering the outcomes of this review, these dimensions seem promising areas of further investigation.

## 5.5 Limitations

As with all research, this study is subject to a number of limitations. Firstly, it should be noted that while the review suggests that there are patterns to be found in the research literature, the analyses made must be interpreted cautiously given that individual studies also caution careful readings of the empirical results. Two main problems were commonly found across a number of included studies. The first problem was that the experimental studies had to generally resort to different levels of semi-randomized designs (rather than full randomization). The second problem was that the number of participants was generally considered to be low. A

suggestion commonly pointed out by researchers is that future studies should attempt to involve a larger number of participants.

Secondly, even though a systematic review sets out to identify all possible research pertaining to a particular research question (Barajas Eriksson et al., 2013), the literature on the more general connections between game-based learning and language learning is simply too vast to include in a single review. It is important to note that there is plenty of research exploring this relationship that was not included in this study, and that the research included pertains to the particular dimensions of language use that is presented in the introduction and method chapters. Particularly, this chosen angle meant that the study did not in any greater length analyze research on commercial games such as MMORPGs, or the differences between first- or second language learning. Thus, these are angles that may be worth exploring further in future studies.

## 6 Conclusion

The results of this review indicate that while there is a number of problems associated with video game-based learning, the included studies indicate that particularly serious games designed to target the most fundamental literacy skills of a person's first language may be used as an effective tool to practice these skills. As a matter of fact, experimental results indicate that this may be a more effective way of learning these skills than through other means. Since these are critical skills for eventual first language fluency, there may be long-term benefits from playing these games, whether that's inside or outside school. Of particular interest is the fact that it seems that many of these games can be used with little or no assistance, and still retain their effect. This can allow teachers to focus on other aspects of teaching, or even compensate (to a degree) for when the pedagogical environment is not ideal.

When it comes to what children older than that can learn from playing video games, studies vary in their approaches both with regard to theory, design and results. Thus, many questions persist. For example, how can serious games employing storytelling, narrative techniques or visual elements increase children's and students' morphological, semantical or pragmatical learning? And how do you effectively foster vocabulary learning through games? From a design perspective, how can designers of serious games move beyond simple multi-choice design to foster more advanced language learning than the most rudimentary literacy skills? Commercial games are certainly of great interest for "older" gamers, but the question of how to research their perceived effectiveness, or how to successfully implement them in a pedagogical setting, remains largely unanswered.

Perhaps the greatest strength of game-based learning, pointed out by a number of researchers, is its *potential* to adapt to the learners' current language level, making sure the player stays within the players' zone of proximal development, to use Vygotsky's classic terminology, as well as to provide sufficient repetitions on learning bottlenecks, which is something that games such as GraphoGame seems to be able to do. However, this may also be an explanation to why certain serious games, such as those that are designed around narratives, seem less successful since they fail to analyze the learner's current language level properly and adapt its content and feedback accordingly.

## 7 References

Articles marked with an \* were included in the review

Abdul Jabbar, A., & Felicia, P. (2015). Gameplay engagement and learning in game-based learning: A systematic review. *Review of educational research*, 85, 740–779. <https://doi.org/10.3102/0034654315577210>

Adams, M. J. (1990). *Beginning to read: Thinking and learning about print*. Massachusetts Institute of Technology.

\*Amorim, A., N., Jeon, L., Abel, Y., Albuquerque, E. X. S., Soares, M., Silva, V. C., Oliveira Neto, J. R. (2022). Escribo play learning games can foster early reading and writing for low-income kindergarten children. *Computers and Education*, 177(2022), 1-13. <https://doi.org/10.1016/j.compedu.2021.104364>

Annetta, L. A., Minogue, J., Holmes, S. Y., & Chen, M. T. (2009). Investigating the impact of video games on high school students' engagement and learning about genetics. *Computers & Education*, 53, 74–85. <https://doi.org/10.1016/j.compedu.2008.12.020>

Association for UK Interactive Entertainment (2022). UK video games fact sheet. London, UK. *2021 key facts about the European video games sector – ISFE*. Accessed on 2022-08-28 from <https://www.isfe.eu/data-key-facts/2021-key-facts-about-europe-s-video-games-sector/>.

Atkinson, T., & Hirumi, A. (2010). The game brain. In Hirumi, A. (Ed.), *Playing games in school* (pp. 57–73). International Society for Technology Education (ISTE).

\*Azman, H. & Dollsaid, N. F. (2018). Applying massively multiplayer online games (MMOGs) in EFL teaching. *Arab World English Journal*, 9(4), 183-18. <https://dx.doi.org/10.24093/awej/vol9no4.1>

\*Baker, D. L., Basaraba, D. L., Smolkowski, K., Conry, J., Hautala, J., Richardson, U., English, S. & Cole, R. (2017). Exploring the cross-linguistic transfer of reading skills in Spanish to English in the context of a computer adaptive reading intervention. *Bilingual Research Journal*, 40(2), 222-239. <http://dx.doi.org/10.1080/15235882.2017.1309719>

Barmark, M. & Djurfeldt, G. (2015). *Statistisk verktygslåda 0: att förstå och förändra världen med siffror*. (1. ed). Studentlitteratur.

Baron-Cohen, S. (2006). The hyper-systemizing, assortative mating theory of autism. *Progress in Neuropsychopharmacology & Biological Psychiatry*, 30, 865–872. doi: 10.1016/j.pnpbp.2006.01.010

\*Bertoni S., Franceschini, S., Puccio, G., Mancarella, M., Gori, S. & Facchetti, A. (2021). Action video games enhance attentional control and phonological decoding in children with developmental dyslexia. *Brain Sciences*, 11(2), 1-18.

<https://doi.org/10.3390/brainsci11020171>

Bielenberg, D., R. & Carpenter-Smith, T. (1997) Efficacy of story in multimedia training. *Journal of Network and Computer Applications*, 20(2), 151–159. doi: 10.1006/jnca.1997.0043

Blair, C. & Razza, R. P. (2007) Relating effortful control, executive function, and false belief understanding to emerging math and literacy ability in kindergarten. *Child Development*, 78, 647–663.

Boy, J., Rensink, R.A., Bertini E., & Fekete, J. D. (2014). A principled way of assessing visualization literacy. *IEEE Transactions on Visualization and Computer Graphics*, 20(12), 10. doi:10.1109/TVCG.2014.2346984

Bowman, R. F. (1982). A Pac-Man theory of motivation. Tactical implications for classroom instruction. *Educational Technology*, 22(9), 14–17.

Buday, B. (2015). Games for health: An opinion. *Games for Health Journal*, 4(1), 38-42. doi: 10.1089/g4h.2014.0083

Cain, K., Catts, H., Hogan, T. & Lomax, R. (2015). Learning to read: Should we keep things simple? *Reading Research Quarterly*, 50, 151–169.

<https://doi.org/10.1002/rrq.99>

\*Calvo-Ferrer, J. R. (2020). Exploring digital nativeness as a predictor of digital game-based L2 vocabulary acquisition. *Interactive Learning Environments*, 28(7), 902-914. <https://doi.org/10.1080/10494820.2018.1548489>

Carlisle, J. F. (1995). Morphological awareness and early reading achievement. In L. B. Feldman (Ed.), *Morphological aspects of language processing* (pp. 189–209). Erlbaum.

Carlisle, J. F. (2000). Awareness of the structure and meaning of morphologically complex words: Impact on reading. *Reading and Writing: An Interdisciplinary Journal*, 12(2000), 169–190. <https://doi.org/doi:10.1023/A:1008131926604>

Carlisle, J. F. & Kearns, D. M. (2017). Learning to read morphologically complex words. In Cain, K., Compton, D. L. & Paurilla, R. K. (Eds.), *Theories of reading*

*development: Studies in written language and literacy* (pp. 191–214). John Benjamins. <https://doi.org/10.1075/swll.15.11car>

Cartwright, K. B., Coppage, E. A., Lane, A. B., Singleton, T., Marshall, T. R. & Bentivegna, C. (2017). Cognitive flexibility deficits in children with specific reading comprehension difficulties. *Contemporary Educational Psychology*, 50, 33–44

Cartwright, K. B., Marshall, T. R., Huemer, C. M. & Payne, J. B. (2019) Executive function in the classroom: cognitive flexibility supports reading fluency for typical readers and teacher-identified low-achieving readers. *Research in Developmental Disabilities*, 88, 42–52.

Cassidy, A.M. (2013). The socio-emotional needs of children with dyslexia in different educational settings in Ireland. *Journal of Research in Special Educational Needs*, 13(1), 79–91.

Chen, Z. (2006). *The effects of multimedia annotations on L2 vocabulary immediate recall and reading comprehension: A comparative study of text-picture and audio-picture annotations under incidental and intentional learning conditions*. [doctoral dissertation, University of South Florida]. <http://scholarcommons.usf.edu/etd/2478>

Chevalier, F., Riche, N. H., Alper, B., Plaisant, C., Boy, J. & Elmqvist, N. (2018) Observations and reflections on visualization literacy in elementary school. *IEEE Computer Graphics and Applications*, 38(3), 21–29. doi: 10.1109/mcg.2018.032421650

Cohen, L., Lawrence, M. & Morrison, K. (2018). *Research methods in education. (Eighth edition)*. Routledge.

Colè, P., Duncan, L. G. & Blaye, A. (2014) Cognitive flexibility predicts early reading skills. *Frontiers in Psychology*, 5, 565

Csikszentmihalyi, M. (1996). *Creativity: Flow and the psychology of discovery and invention*. Harper Collins.

DeWalt, D. A., Berkman, N. D., Sheridan, S., Lohr, K. N., & Pignone, M. P. (2004). Literacy and health outcomes. *Journal of General Internal Medicine*, 19(12), 1228–1239.

\*Dore, R. A., Shirilla, M., Hopkins, E., Collins, M., Scott, M., Schatz, J., Lawson-Adams, J., Valladares, T., Foster, L., Puttre, H., Toub, T.S., Hadley, E., Golinkoff, R.M., Dickinson, D. & Hirsh-Pasek, K. (2019). Education in the app store: using a mobile game to support U.S. preschoolers' vocabulary learning. *Journal of Children and Media*, 13(4), 452-471. <https://doi.org/10.1080/17482798.2019.1650788>

- Dornyei, Z. & Ushioda, E. (2009). *Motivation, language, identity and the L2 self*. Multilingual Matters.
- Ehri, L. C., Nunes, S. R., Willows, D. M., Schuster, B. V., Yaghoub-Zadeh, Z. & Shanahan, T. (2001). Phonemic awareness instruction helps children learn to read: Evidence from the national reading panel's meta-analysis. *Reading Research Quarterly*, 36(3), 250–287. <https://doi.org/10.1598/RRQ.36.3.2>
- Ehri, L. C., Nunes, S. R., Stahl, S. & Willows, D. M. (2001). Systematic phonics instruction helps students learn to read: evidence from the national reading panel's meta-analysis. *Reading Research Quarterly*, 36(3), 250–287. <https://doi.org/10.3102/00346543071003393>
- Ehri, L. (2005). Learning to read word: theory, findings, and issues. *Scientific Studies of Reading*, 9, 167–188. doi:10.1207/s1532799xssr0902\_4
- Eichenbaum, A., Bavelier, D. & Green, C. S. (2014). Video games: play that can do serious good. *American Journal of Play*, 7(1), 50–61.
- \*Ensmann, S. (2021). Digital games to improve learning in Haiti. *TechTrends: Linking Research and Practice to Improve Learning*, 65(5), 884-895. <http://dx.doi.org/10.1007/s11528-021-00630-8>.
- Eriksson Barajas, K., Forsberg, C & Wengström, Y. (2013). *Systematiska litteraturstudier i utbildningsvetenskap: vägledning vid examensarbeten och vetenskapliga artiklar* (1 ed.). Natur & Kultur.
- Felton, M., Crowell, A., & Liu, T. (2015). Arguing to agree: Mitigating my-side bias through consensus-seeking dialogue. *Written Communication*, 32(3), 317–331.
- Fisch, S. M. (2016). The Capacity Model, 2.0: cognitive processing in children's comprehension of educational games. In J. T. Piotrowski (Ed.), *Digital media and early literacy. Symposium conducted at the society for research on child development special topics meeting*. Technology and Media in Children's Development.
- Franceschini, S., Gori, S., Rufno, M., Pedrolli, K. & Facoetti, A. (2012). A causal link between visual spatial attention and reading acquisition. *Current Biology* 22, 814–819.
- Frost, J., Madsbjerg, S., Niedersøe, J., Olofsson, Å & Sørensen, P. M. (2005). Semantic and phonological skills in predicting reading development from 3-16 years of age. *Dyslexia*, Vol. 11 (2), 79-92.
- Galuschka, K., Ise, E., Krick, K. & Schulte-Körne, G. (2014). Effectiveness of treatment approaches for children and adolescents with reading disabilities: A meta-analysis of randomized controlled trials. *PLoS ONE*, 9(2), e89900. <https://doi.org/10.1371/journal.pone.0089900>

- Garris, R., Ahlers, R., & Driskell, J. (2002). Games, motivation, and learning: A research and practice model. *Simulation & Gaming, 33*, 441–467. <https://doi.org/10.1177/1046878102238607>
- Gee, J. (2003) What video games have to teach us about learning and literacy. *Computers in Entertainment, 1*:20. doi: 10.1145/950566.950595
- Gee, J. (2007). *What video games have to teach us about learning and literacy* (2nd ed.) Palgrave MacMillan.
- Gee, J. (2008). Video games and embodiment. *Games and Culture, 3*(3-4), 253–263. <https://doi.org/10.1177/1555412008317309>
- Gee, J. (2010a). New digital media and learning as an emerging area and “worked examples” as one way forward. *The John D. and Catherine T. MacArthur Foundation Series on Digital Media and Learning*. The MIT Press.
- Gee, J. (2010b). Video games: What they can teach us about audience engagement. *Nieman Reports, 64*(2), 52–54. Retrieved from Communication & Mass Media Complete Database
- Gee, J. (2015). Discourse analysis of games. In R. H. Jones, Chik, A. & Hafner, C. A. (Eds.), *Discourse and digital practices: Doing discourse analysis in the digital age* (pp. 18–27). Routledge.
- Graff, G. (2004). *Clueless in academe: How schooling obscures the life of the mind*. Yale University Press.
- Gough, P. B., & Tunmer, W. E. (1986). Decoding, reading, and reading disability. *Remedial and Special Education (RASE), 7*, 6–10. doi:10.1177/074193258600700104
- Grainger, J., Tydgate, I. & Issele, J. (2010) Crowding affects letters and symbols differently. *Journal of Experimental Psychology: Human Perception and Performance, 36*, 673–688
- Grainger, J., Dufau, S. & Ziegler, J. C. (2016) A vision of reading. *Trends in Cognitive Sciences, 20*, 171–179
- Grant, A., Wood, E., Gottardo, A., Evans, M. A., Phillips, L., & Savage, R. (2012). Assessing the content and quality of commercially available reading software programs: Do they have the fundamental structures to promote the development of early reading skills in children? *NHSA Dialog, 15*, (319e342). <https://doi.org/10.1080/15240754.2012.725487>

- Grigorenko, E. L., Klin, A., & Volkmar, F. (2003). Annotation: hyperlexia: disability or superability? *Journal of Child Psychology and Psychiatry*, *44*, 1079–1091. doi: 10.1111/1469-7610.00193
- Gunter, G. A., Kenny, R. F., & Vick, E. H. (2008). Taking educational games seriously: Using the RETAIN model to design endogenous fantasy into standalone educational games. *Educational Technology Research and Development*, *56*(5–6), 511–537
- Gustafson, K. L., & Branch, R. M. (2002). *Survey of instructional design models* (Fourth ed.). Eric Clearinghouse on Information and Technology
- Ham, B., Stoolmiller, M. & Chard, D. J. (2015). Measuring the dimensions of alphabetic principle on the reading development of first graders: the role of automaticity and unitization. *Journal of Learning Disabilities*, *41*(2), 143–157. <https://doi.org/10.1177/0022219407313585>
- Hirsh-Pasek, K., Zosh, J. M., Golinkoff, R. M., Gray, J. H., Robb, M. B., & Kaufman, J. (2015). Putting education in “educational” apps: Lessons from the science of learning. *Psychological Science in the Public Interest*, *16*(1), 3–34.
- \*Hobbs, L., Hartley, C., Bentley, S., Bibby, J., Bowden, L., Hartley, J. & Stevens, C. (2020). Shared special interest play in a specific extra-curricular group setting: A Minecraft club for children with special educational needs. *Educational & Child Psychology*, *34*(4), 81-95.
- Huizenga, J., Admiraal, W., Akkerman, S., & ten Dam, G. (2009). Mobile game-based learning in secondary education: Engagement, motivation and learning in a mobile city game. *Journal of Computer Assisted Learning*, *25*(2009), 332–344. <https://doi.org/10.1111/j.1365-2729.2009.00316.x>
- Hung, H., Yang, J. & Hwang, G. (2018). A scoping review of research on digital game-based language learning. *Computers and Education*, *126*, 89-104.
- \*Huynh, E., Nyhout, A., Ganea, P. & Chevalier, F. (2021). Designing narrative-focused role-playing games for visualization literacy in young children. *IEEE Transactions on Visualization and Computer Graphics*, *27*(2), 924-934.
- Justice, L. J., & Ritzhaupt, A. D. (2015). Identifying the barriers to games and simulations in education creating a valid and reliable survey. *Journal of Educational Technology Systems*, *44*(1), 86–125.
- Kamio, Y., & Toichi, M. (2000). Dual access to semantics in autism: is pictorial access superior to verbal access? *Journal of Child Psychology and Psychiatry*, *41*, 859–867. doi: 10.1111/1469-7610.00673

- Kana, R. K., Keller, T. A., Cherkassky, V. L., Minshew, N. J., & Just, M. A. (2006). Sentence comprehension in autism: thinking in pictures with decreased functional connectivity. *Brain: a Journal of Neurology*, *129*, 2484–2493. doi: 10.1093/brain/awl164
- Kebritchi, M., & Hirumi, A. (2008). Examining the pedagogical foundations of modern educational computer games. *Computers & Education*, *51*(4), 1729–1743. <https://doi.org/10.1016/j.compedu.2008.05.004>
- Ke, F., & Abras, T. (2013). Games for engaged learning of middle school children with special learning needs. *British Journal of Educational Technology*, *44*, 225–242. <https://doi.org/10.1111/j.1467-8535.2012.01326.x>
- Kenny, R. F. (2011). Beyond the Gutenberg parenthesis: Exploring new paradigms in media learning. *Journal of Media Literacy Education*, *3*(1), 32–46.
- Kiili, K. (2005). Digital game-based learning: Towards an experiential gaming model. *Internet and Higher Education*, *8*, 13–24. <https://doi.org/10.1016/j.iheduc.2004.12.001>
- Kim, S., Chang, M., Deater-Deckard, K., Evans, M. A., Norton, A., & Samur, Y. (2017). Educational games and students' game engagement in elementary school classrooms. *Journal of Computers in Education*, *4*, 395–418. <https://doi.org/10.1007/s40692-017-0095-4>
- Koshino, H., Carpenter, P. A., Minshew, N. J., Cherkassky, V. L., Keller, T. A. & Just, M. A. (2005). Functional connectivity in an fMRI working memory task in high-functioning autism. *Neuroimage*, *24*, 810–821. doi: 10.1016/j.neuroimage.2004.09.028
- \*Lawrence, A. & Sherry, M. (2021). How feedback from an online video game teaches argument writing for environmental action. *Journal of literacy research*, *53*, 29-52. <http://dx.doi.org/10.1177/1086296X20986598>
- Lepper, M. & Malone, T. (1987). Intrinsic motivation and instructional effectiveness in computer based education. In R. E. Snow & M. J. Farr (Eds.), *Aptitude, learning and instruction: III. Cognitive and affective process analyses* (pp. 223–253). Erlbaum.
- Lieberman, D. A. (2006) What can we learn from playing interactive games? In P. Vorderer & J. Bryant (Eds.), *Playing video games: Motives, responses, and consequences* (pp. 379–397). Lawrence Erlbaum Associates Publishers.
- Lieberman, D., Bates, C. & So, J. (2009) Young children's learning with digital media. *Computers in the schools*, *26*(4), 271-283. doi: 10.1080/07380560903360194

- Linderoth, J. (2012). Why gamers don't learn more: An ecological approach to games as learning environments. *Journal of gaming & virtual worlds* 4(1), 45-62. doi:10.1386/jgvw.4.1.45\_1
- Lundberg, I., Frost, J., & Peterson, O. (1988). Effects of an extensive program for stimulating phonological awareness in preschool children. *Reading Research Quarterly*, 13(3), 263–284.
- Lundberg, I. (2010). *Läsningens psykologi och pedagogik*. Natur & Kultur.
- Lyytinen, H., Semrud-Clikeman, M. & Li, H. (2021). Supporting acquisition of spelling skills in different orthographies using an empirically validated digital learning environment. *Frontiers in Psychology* 12(566220). doi: 10.3389/fpsyg.2021.566220
- McCarrick, K. & Li, X. (2007). Buried treasure: The impact of computer use on young children's social, cognitive, language development and motivation. *AACE Journal*, 15, 73–95.
- McLaughlin, M. J., Speirs, K. E., & Shenassa, E. D. (2014). Reading disability and adult attained education and income: evidence from a 30-year longitudinal study of a population-based sample. *Journal of Learning Disabilities*, 47(4), 374–386.
- Mercer, N. & Littleton, K. (2007). *Dialogue and the development of children's thinking: A sociocultural approach*. Routledge.
- Miller, L. M., Chang, C., Wang, S., Beier, M. E., & Klisch, Y. (2011). Learning and motivational impacts of a multimedia science game. *Computers & Education*, 57(2011), 1425–1433. <https://doi.org/10.1016/j.compedu.2011.01.016>
- Mills, K. (2015). *Literacy theories for the digital age: social, critical, multimodal, spatial, material and sensory lenses*. Multilingual matters.
- Newell, G. E., Beach, R., Smith, J. & VanDerHeide, J. (2011). Teaching and learning argumentative reading and writing: A review of research. *Reading Research Quarterly*, 46(3), 273–304.
- Noldus, L. P., Spink, A. J., & Tegelenbosch, R. A. (2001). EthoVision: A versatile video tracking system for automation of behavioral experiments. *Behavior Research Methods, Instruments & Computers*, 33(3), 398–414.
- Osguthorpe, R. T., & Osguthorpe, L. S. (2009). *Choose to learn. Teaching for success every day*. Corwin Press.
- O'Sullivan, M., Robb, N., Howell, S., Marshall, K. & Goodman, L. (2017). Designing inclusive learning for twice exceptional students in Minecraft. *International Journal of E-Learning and Distance Education*, 32(2), 1–25.

- Oxman, S., & Wong, W. (2014). *White paper: Adaptive learning systems*. Integrated Education Solutions.
- Paas, F., Renkl, A., & Sweller, J. (2003). Cognitive load theory and instructional design: recent developments. *Educational Psychologist*, 38(1), 1–4.
- \*Pasqualotto, A., Altarelli, I., De Angeli, A., Menestrina, Z., Bavelier, D. & Venuti, P. (2022). Enhancing reading skills through a video game mixing action mechanics and cognitive training. *Nature Human Behaviour*, 6, 545–554.  
<https://doi.org/10.1038/s41562-021-01254-x>
- Pennington, B. F., & Lefly, D. L. (2001). Early reading development in children at family risk for dyslexia. *Child Development*, 72(3), 816–833.
- Perfetti, C. A. & Stafura, J. (2014). Word knowledge in a theory of reading comprehension. *Scientific Studies of Reading*, 18, (1), 22–37.  
<https://doi.org/10.1080/10888438.2013>
- Prensky, M. (2001). *Digital game-based learning*. McGraw-Hill.
- Prensky, M. (2010). Educating the millennial generation. In Hirumi, A. (Ed.), *Playing games in school* (pp. 7–32). International Society for Technology in Education (ISTE).
- Peng, J. E. (2014). *Willingness to Communicate in the Chinese EFL University Classroom: An Ecological Perspective*. Multilingual Matters.
- \*Rankin, Y. A., Tibi, S., Kennington, C. & Han, N. (2021). In-game social interactions to facilitate ESL students’ morphological awareness, language and literacy skills. *Proceedings of the ACM on Human Computer Interaction*, 5(Chi Play).  
<https://doi.org/10.1145/3474706>
- Reinders, H. & Wattana, S. (2012). Talk to me! Games and students’ willingness to communicate. In Reinders, H. (ed.) *Digital games in language learning and teaching. New language learning and teaching environments* (pp. 156-187) Palgrave Macmillan. [https://doi.org/10.1057/9781137005267\\_9](https://doi.org/10.1057/9781137005267_9)
- Ringland, K. E., Wolf, C. T., Faucett, H., Dombrowski, L. & Hayes, G. R. (2016). ‘Will I always be not social?’: Re-conceptualising sociality in the context of a *Minecraft* community for autism. *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, 1256– 1269.
- Ronimus, M., Kujala, J., Tolvanen, A., & Lyytinen, H. (2014). Children’s engagement during digital game-based learning of reading: The effects of time, rewards, and challenge. *Computers & Education*, 71, 237–246.  
<https://doi.org/10.1016/j.compedu.2013.10.008>

- \*Ronimus, M., Eklund, K., Pesu, L. & Lyytinen, H. (2019). Supporting struggling readers with digital game-based learning. *Educational Technology Research and Development*, 67(3), 639-663. <https://doi.org/10.1007/s11423-019-09658-3>
- Rosman, F., Alias, N. & Siraj, S. (2013). Potential of video games in learning Bahasa Melayu vocabulary among international university students in Malaysia: A meta analysis of selected journals. *Turkish Online Journal of Educational Technology* 12(4), 163-170.
- Ruz, M. & Nobre, A. C. (2008) Attention modulates initial stages of visual word processing. *Journal of Cognitive Neuroscience*,. 20, 1727–1736
- Samur, Y. (2012). *Measuring engagement effects of educational games and virtual manipulatives on mathematics* [Doctoral Dissertation, Virginia Polytechnic Institute and State University].
- \*Samur, Y. (2019). Kes Sesi: A mobile game designed to improve kindergarteners' recognition of letter sounds. *Journal of Computer Assisted Learning*, 35(2), 294–304. <https://doi.org/10.1111/jcal.12331>
- \*Sánchez-Vincitore, L. & Cross, A. (2021). Effects of an electronic game on early literacy skills. *Psicologia Escolar e Educacional*, 25. <http://dx.doi.org/10.1590/2175-35392021225567>
- Schell, J. (2008). *The art of game design: A book of lenses*. Morgan Kaufmann Publishers. <https://doi.org/10.1201/9780080919171>
- \*Schmitt, K., Hurwitz, L., Sheridan, D. & Nichols Linebarger, D. L. (2018). Learning through play: The impact of web-based games on early literacy development. *Computers in Human Behaviour* 81(2018), 378–389. <https://doi.org/10.1016/j.chb.2017.12.036>
- Schneider, E. F. (2004) Death with a story. *Human Communication Research*, 30(3), 361–375. doi:10.1111/j.1468-2958.2004.tb00736.x
- \*Serra, J. & Gilabert, R. (2021). Algorithmic versus teacher-led sequencing in a digital serious game and the development of second-language reading fluency and accuracy. *British Journal of Educational Technology* 52(5). 1898-1916. Doi: 10.1111/bjet.13131
- \*Serret, S., Hun, S., Thümmler, S., Pierron, P., Santos, A., Bourgeois, J. & Askenazy, F. (2017). Teaching literacy skills to French minimally verbal school-aged children with autism spectrum disorders with the serious game SEMA-TIC: An exploratory study. *Frontiers in Psychology*, 8(SEP). doi: 10.3389/fpsyg.2017.01523

Shute, V., Ventura, M., & Bauer, M. (2009). Melding the power of serious games and embedded assessment to monitor and foster learning: Flow and grow. In Ritterfeld, U., Cody, M & Vorderer, P. (Eds.), *Serious games: Mechanisms and effects* (pp. 295–321). Routledge.

Squire, K. (2003). Video games and education. *International Journal of Intelligent Simulations and Gaming*, 2, 49

Stahl, G. (2006). *Group cognition: Computer support for building collaborative knowledge*. MIT Press.

Strijkers, K., Bertrand, D. & Grainger, J. (2015) Seeing the same words differently: the time course of automaticity and top-down intention in reading. *Journal of Cognitive Neuroscience*, 27, 1542–1551

Takac, V. P. (2008). *Vocabulary learning strategies and foreign language acquisition*. Multilingual Matters.

Taylor, C. (2003) *New kinds of literacy, and the world of visual information*. Literacy.

Taylor, J. S. H., Davis, M. H. & Rastle, K. (2017). Comparing and validating methods of reading instruction using behavioural and neural findings in an artificial orthography. *Journal of Experimental Psychology: General*, 146(6), 826. <http://dx.doi.org/10.1037/xge0000301>

Teo, T. (2013). An initial development and validation of a Digital Natives Assessment Scale (DNAS). *Computers & Education*, 67, 51–57.

\*Thomson, J. M., Foldnes, N., Uppstad, P. H., Njå, M., Solheim, O. J. & Lundetr, K. (2020) Can children’s instructional gameplay activity be used as a predictive indicator of reading skills? *Learning and Instruction*, 68. 10.1016/j.learninstruc.2020.101348

Tuedor, M. (2006). Universal access through accessible computer educational programs to develop the reading skills of children with autistic spectrum disorders. *Universal Access in the Information Society*, 5, 292–298. doi: 10.1007/s10209-006-0047-7

Vandewaetere, M., Desmet, P. & Clarebout, G. (2011). The contribution of learner characteristics in the development of computer-based adaptive learning environments. *Computers in Human Behavior*, 27(1), 118–130.

<https://doi.org/10.1016/j.chb.2010.07.038>

- \*Vasileiadou, I. & Makrina, Z. (2017). Using online computer games in the ELT classroom: A case study. *English language teaching* 10(12), 134–150.  
<http://doi.org/10.5539/elt.v10n12p134>
- \*Vasquez-Calvo, B. (2018) The online ecology of literacy and language practices of a gamer. *Educational Technology & Society*, 21(3), 199–212.
- Vygotsky, L. S. (1930-1934/1978). *Mind in society: The development of higher psychological processes*. Harvard University Press
- Wagner, R. K., Torgesen, J. K., Rashotte, C. A., Hecht, S. A., Barker, T. A., Burgess, S. R., Donahue, J. & Garon, T. (1997). Changing relations between phonological processing abilities and word-level reading as children develop from beginning to skilled readers: A 5- year longitudinal study. *Developmental Psychology*, 33, 468–479.
- Wang, H. & Singhal, A. (2009). Entertainment-education through digital games. In Ritterfield, U., Cody, M. & Vorderer, P. (eds.), *Serious games: mechanisms and effects* (pp. 271–292). Routledge
- Wartella, E., Beaudoin-Ryan, L., Blackwell, C. K., Cingel, D. P., Hurwitz, L. B., & Lauricella, A. R. (2016). What kind of adults will our children become? The impact of growing up in a media-saturated world. *Journal of Children and Media*, 10, 13-20. <https://doi.org/10.1080/17482798.2015.1124796>
- Weston, C., McAlpine, L., & Bordonaro, T. (1995). A model for understanding formative evaluation in instructional design. *Educational Technology Research and Development*, 43(3), 29–48
- Williams, C., Wright, B., Callaghan, G., & Coughlan, B. (2002). Do children with autism learn to read more readily by computer assisted instruction or traditional book methods? A pilot study. *Autism: The International Journal of Research and Practice*, 6, 71–91. doi: 10.1177/1362361302006001006
- Wolfswinkel, J., Furtmueller, E., & Wilderom, C. (2013). Using grounded theory as a method for rigorously reviewing literature. *European Journal of Information Systems*, 22(1), 45–55.
- Wouters, P., van Nimwegen, C., van Oostendorp, H., & van der Spek, E. D. (2013). A metaanalysis of the cognitive and motivational effects of serious games. *Journal of Educational Psychology*, 105(2), 249–265.
- Wrzesien, M., & Raya, M. A. (2010). Learning in serious virtual worlds: Evaluation of learning effectiveness and appeal to students in the E-junior project. *Computers & Education*, 55, 178–187. <https://doi.org/10.1016/j.compedu.2010.01.003>

Yelland, N. & Masters, J. (2007). Rethinking scaffolding in the information age. *Computers & Education*, 48, 362–382. <https://doi.org/10.1016/j.compedu.2005.01.010>

Xu, Z., Chen, Z. & Eutsler, L. (2020). A scoping review of digital game-based technology on English language learning. *Educational Technology Research and Development*, 68(3), 877–904. <https://doi.org/10.1007/s11423-019-09702-2>

# 8 Appendix

## 8.1 Glossary

The purpose of the following section is to provide an overview of terms and concepts occurring in the included articles or that are otherwise utilized in this essay. While there may be many interpretations of a term, this is how they are understood in this study. By necessity these descriptions are very brief; they are not by any means intended to be conclusive. Theories on game-based learning, literacy or language learning, as they are understood by researchers presented in the essay, are elaborated upon further in the results section. Commonly used concepts are also included (without reference) to clarify my understanding of these terms.

*Commercial games:* Video games widely available, usually designed for entertainment purposes. Designed by video game companies within the video games industry rather than people within academics (Buday, 2015)

*Decoding:* The ability to correctly identify individual graphemes and correctly associate them with corresponding phonemes (Lundberg, 2010)

*DGBLL:* An abbreviation meaning digital game-based language learning. The term can be used to point out the distinction between digital and physical games

*Game-based learning:* A commonly used term within the field which implies that learning can be facilitated through playing different kinds of games

*Grapheme:* A graphical representation of a language sound in the form of a written symbol, usually a letter or a combination of letters (Lundberg, 2010)

*Literacy:* Strategies related to reading skills, it is now a concept used to describe a number of practices related to cognitive skills needed for “reading” of physical and digital artefacts. In this study, the term *fundamental literacy skills* (Schmitt et al., 2018), or *pre-reading skills*, is used to describe those aspects of literacy, such as phonological awareness and alphabet and letter knowledge, which are a prerequisite for fluency of reading and writing. These skills are often learnt together around the time of school entry. In contrast to this, other strands of literacy research exist such as *visual literacy* (Huynh et al., 2021), *argument literacy* (Lawrence & Sherry, 2021) or *New literacy studies* (Mills, 2015, Vasquez-Calvo, 2018), illustrating the complexity of the term. These schools of thought generally suggest that “reading” of different

artefacts requires a wider scope of skills than what has traditionally been suggested (Mills, 2015)

*MMORPG*: A widely used abbreviation for *Massively multiplayer online role-playing game*. A game largely built around social dynamics with a vast number of simultaneous participants online

*Morpheme*: These easily recognizable “building blocks” of a language are called *morphemes*, and they are the smallest meaning-making units of a language. For instance, while the word “table” consists only of what is called the stem, the word “tables” consists of this stem with the addition of the morpheme “s”. This small but significant grammatical addition to the word changes the semantic meaning of the word entirely, from indicating a single object to indicating the plural form of the word (Lundberg, 2010)

*Morphological awareness*: Being able to understand that the meaning of words can be changed by manipulating grammatical forms, for example by adding suffixes such as an -s to get the plural form (such as changing the word *table* into *tables*). Carlisle considers it a critical skill for second language vocabulary acquisition (Carlisle, 1995)

*Phonological awareness*: The understanding that languages consists of different language sounds, or phonemes, and that each phoneme has a corresponding *grapheme* used to represent it in writing. In other words, correctly grasping “the alphabetic principle”. Considered a critical skill for learning to read and highly predictive of later academic performance (Frost et al., 2005) along with closely related abilities such as word decoding and letter knowledge (Pennington & Lefly, 2001)

*Phoneme*: A language sound, the smallest identifiable sound unit of a language (Lundberg, 2010)

*Phonemic awareness*: The ability to properly distinguish between different language sounds, or phonemes (Lundberg, 2010)

*Pragmatic awareness*: The dimension of language use related to situational awareness, for example being able to adjust one’s language use depending on genre when writing, or following social rules regarding turn-taking when communicating with another person (Lundberg, 2010)

*Semantics*: The actual meaning (or idea) represented by a word. Related to our ability to correctly represent and interpret information, and to understand how different ideas relate to each other (Lundberg, 2010)

*Semantic awareness*: Understanding the relationship between words and the reality that they refer to (Lundberg, 2010)

*Serious games*: A term widely used to describe games designed specifically for learning purposes. Often designed by researchers or educational specialists rather than commercially driven companies.

*The alphabetic principle:* The understanding that letters (so called graphemes) represent language sounds (phonemes), and that letter combinations represent different sounds, is often called the *alphabetic principle* (Adams, 1990)

*The Simple View of Reading:* This theory can be expressed in the form of the equation “reading = decoding x understanding” and it points out that reading is dependent on both technical skills (being able to correctly associate graphemes with phonemes) but also being able to understand the semantical meaning behind words. Without both components, reading essentially fails. The theory has had great impact on research on children’s fundamental literacy skills (Gough & Tunmer, 1986)

*Video game:* In this review, the commonly used term video game is used to indicate games played with a screen, for example using a mobile phone, a computer or a games console. This differentiates these types of games from games consisting of physical components, such as board games. Video games can in turn be categorized into commercial games and serious games (see respective section in glossary)

## 8.2 Overview of articles

Table of included articles (N = 21):

Author/Title	Aim/goals of GBL	Research design/method	Participants, total (completing the study)	Relevant language aspects	Main findings
Amorim et al. (2022). <i>Escribo play learning games can foster early reading and writing for low-income kindergarten children</i>	Investigate potential learning gains of cost-effective game based learning and identify potential barriers for deployment	Mixed methods: Experiment (randomised controlled trial) + qualitative data to interpret results. 3 months’ duration, twice a week.	749 pre-schoolers across 62 classrooms in 17 Brazilian private schools of five different cities	Phonological awareness, phonemic awareness, letter-sound correspondence activities	Significant gains in word reading and word writing despite barriers faced

Azman, H. and Dollsaid, N. (2018). <i>Applying massively multiplayer online games (MMORPGs) in EFL teaching</i>	Investigate language learning potential by using MMORPGs from an EFL-perspective	Qualitative case study: observations from 10 game sessions, data from questionnaires, analysis of 10 in-game chat logs and email interviews (content analysis)	Five participants aged 18-21	Inductive	Collaborative exchanges, team spirit, and strategic social competence benefits general language learning Correspondence with a number of Gee's (2003) principles
Baker, D. et al. (2017). <i>Exploring the cross-linguistic transfer of reading skills in Spanish to English in the context of a computer adaptive reading intervention</i>	Increase decoding skills and oral reading fluency across languages	Experiment	78 first grade Spanish speaking children across five classrooms in Texas	Phonemic awareness, letter-sound correspondence knowledge, word recognition	Mixed results, may be effects across languages, but more studies are needed  Strong student engagement
Bertoni et al. (2021). <i>Action video games enhance attentional control and phonological decoding in children with developmental dyslexia</i>	Investigate whether features of action-oriented video games may improve fundamental literacy skills	Experiment: 12 hours of playing action-oriented video games	14 native Italian children with developmental dyslexia	Phonological decoding, pseudoword reading, word reading	Improvement of pseudoword reading speed (phonological decoding) and attentional control however word reading did not improve
Calvo-Ferrer, R. (2020). <i>Exploring digital nativeness as a predictor of digital game-based L2 vocabulary acquisition</i>	Investigate whether level of digital nativeness has bearing on vocabulary learning outcomes	Quasi-experiment: Questionnaires, knowledge tests, digital native assessment	86 university students learning English as a foreign language	Second language vocabulary acquisition	No major differences between students, though some indications that some students may be more distracted by gaming elements than others
Dore et al. (2019).	Verify claims made in learning apps, provide researched	2 experiments	Study 1: 57 girls aged 4	Vocabulary	Significant gains for target words,

<i>Education in the app store: using a mobile game to support U.S. preschoolers' vocabulary learning.</i>	games as an alternative to commercial products for fundamental literacy skills		Study 2: 33 children aged 3-4		also with regard to children's using the words in their own production.  Positive results even outside experimental setting.
Ensmann, S. (2021). <i>Digital games to improve learning in Haiti</i>	Promote English-language learning among disadvantaged Haitian children	Case study	23 Children with Creole as native language and low-level of English	Semantics/vocabulary but also letters of the alphabet	Important design implications: reliability, cues/hints, clear navigational elements
Hobbs et al. (2020). <i>Shared special interest play in a specific extra-curricular group setting: A Minecraft club for children with special educational need</i>	Investigate language learning opportunities in a game-based extra-curricular setting	Mixed methods: qualitative (interviews) and quantitative (surveys). Data collected during a four year period (2015-2019)	In total 29 children, among them 23 boys, aged 5-14. (this amounted to 26,6 of all club members) 37 caretakers.	Inductive study – no predetermined focus apart from the concept of learning	High degree of belief among children and parents that the activities benefited communication skills such as taking social initiatives (pragmatics)
Huynh et al. (2021). <i>Designing narrative-focused role-playing games for visualization literacy in young children</i>	Investigate whether visualization literacy and narrative elements have any impact on language learning outcomes	Mixed methods: Experiment with randomization (comparison of two versions of game) + interviews	33 children, 11-13 years old	Narration (semantics), Visualization literacy	Results on learning mostly inconclusive. Unclear whether changes in performance between two groups was the result of guessing or the result of improved learning. Mixed reactions and opinions from participants.
Lawrence & Sherry (2021). <i>How feedback from an online</i>	Investigate the pedagogical quality of video game	Qualitative, discourse-analytic case study (recorded screen	114 seventh-graders, 10	Writing argumentation, logical sequencing and semantics	Students written propositions are mostly

<i>video game teaches argument writing for environmental action</i>	feedback on students' argumentative writing	capture, interviews, analysis of students' writing)	focal students		inadequate and not based on semantics that the game provides
Pasqualotto et al. (2022). <i>Enhancing reading skills through a video game mixing action mechanics and cognitive training</i>	Investigate whether a video game-based intervention based on theories of attentional control could improve literacy skills	Experiment (with follow-up test 6 months later)	151 children, aged 8-14	Reading	Reading errors lessened and reading speed improved, attentional skills and planning skills significantly improved  Small but significant correlation with grades in Italian (delayed evaluation)
Rankin et al. (2021). <i>In-game social interactions to facilitate ESL-students morphological awareness, language and literacy skills</i>	Facilitate social interactions between native speakers and second language learners to improve morphological awareness and increase vocabulary learning among the latter group	Quasi-experiment (analysis of chat-log, interviews, focus group discussion)	Four dyads of students (in each pair there was one first language learner and one second language learner)	Morphology, vocabulary and willingness to communicate (WTC) in a second language acquisition (SLA) context	Mixed: Only significant improvement of word reading, in both groups.  Communicative/pragmatic dimensions dominated other aspects. Friendly and attentive social interactions vital for increasing quality of written output
Ronimus et al. (2019). <i>Supporting struggling readers with digital game-based learning</i>	Investigate effectiveness of digital game-based learning for users with moderate to severe learning difficulties	Mixed: Experiment + qualitative (self-report measures, teachers' and parents evaluation). Randomization of schools	37 native speakers of Finnish from 25 schools, mean age 8.23 years	Word reading, reading fluency, reading comprehension, spelling	Improvement of childrens' word reading skill but no transfer to spelling, sentence-level reading fluency or reading comprehension.

					Cognitive engagement crucial for optimal learning
Samur, Y. (2019). <i>Kes Sesi: a mobile game designed to improve kindergarteners' recognition of letter sounds</i>	Improve letter and sound recognition through gameplay	Quasi-experiment	87 kindergarteners, 4-5 year old	Phonological awareness, letter recognition	Significant gains with regard to sound-recognition performance, otherwise mixed
Sanchez-Vincitore & Cross (2021). <i>Effects of an electronic game on early literacy skills</i>	Increasing access and exposure for students from the Dominican Republic to the alphabetic principle to benefit reading automaticity	Experiment, 10 sessions of 20 minutes each	19 students, mean age 5.37	Phonological awareness (The alphabetic principle, decoding) and reading	Significant gains of syllable recognition but not in the case of other reading skills
Schmitt et al. (2018). <i>Learning through play: The impact of web-based games on early literacy development</i>	Improve fundamental literacy skills, particularly with regard to the needs of socially disadvantaged children	Experiment (with randomization)	136 preschoolers	Concepts of print, alphabetic knowledge, phonemic awareness, phonics, fluency, comprehension	Significant gains on several literacy skills, particularly phonological awareness, phonics and vocabulary
Serra & Gilabert (2021). <i>Algorithmic versus teacher-led sequencing in a digital serious game and the development of second language reading fluency and accuracy</i>	Explore the significance of learning sequences and its effect on reading development	Experiment	67 children 11/12 years old, from four intact groups	Phonological awareness/decoding, morphological awareness, syntax, vocabulary	Algorithmic sequencing resulted in significantly stronger improvement of decoding skills for participants with initially weaker decoding abilities
Serret et al. (2017). <i>Teaching literacy skills to French minimally verbal school-aged children with autism spectrum disorders with the serious game SEMA-TIC</i>	Investigate the effectiveness of SEMA-TIC, a game designed based on non-verbal cognitive skills. Whole word reading and visual elements for autistic learners	Experiment (non-randomized)	25 children with autism, 6-11 years	Alphabet knowledge, word reading, word-non-word discrimination, sentence reading, word segmentation	Significant gains in alphabet knowledge. Literacy conventions and decoding skills.

Thomson et al. (2020). <i>Can children's instructional gameplay activity be used as a predictor of reading skills?</i>	Investigate whether the game Grapho-Game can be used to assess whether children's risk status, second language status or gender can predict their learning growth	Experiment (longitudinal study, children played for 25 weeks)	137 six-year old children	Phonological awareness, word level reading	Second language status and gender did not predict significant variation of growth curve parameters, but children's risk status did.
Vasileiadou & Makrina (2017). <i>Using online computer games in the ELT classroom: A case study</i>	Investigate whether computer games make learning English vocabulary more effective and enjoyable	Mixed: Experiment + interviews	33 four-grade students	Vocabulary	Initially lower performing students improved notably more in the experimental group. Students liked learning vocabulary through games and also considered it effective.
Vasquez-Calvo, B. (2018). <i>The online ecology of literacy and language practices of a gamer</i>	Investigate language practices and language learning of a "gamer"	Qualitative, exploratory case study (interviews, online observations, screencast videos)	One Spanish man, aged 30, who had learnt English in school and now learnt more through his gaming practices	Inductive	The participant developed literacy skills (IT, linguistic, sociocultural) in tandem with language skills: reading, writing, semantics, pragmatics, spelling, grammar, lexical-syntax