JYU DISSERTATIONS 210

Amin Rasti Behbahani

Investigating The Effect of Digital Game Tasks, Inducing Different Levels of Involvement Load, on the Acquisition of Vocabulary Items

UNIVERSITY OF JYVÄSKYLÄ
FACULTY OF HUMANITIES AND
SOCIAL SCIENCES

JYU DISSERTATIONS 210

Amin Rasti Behbahani

Investigating the Effect of Digital Game Tasks, Inducing Different Levels of Involvement Load, on The Acquisition of Vocabulary Items

Esitetään Jyväskylän yliopiston humanistis-yhteiskuntatieteellisen tiedekunnan suostumuksella julkisesti tarkastettavaksi yliopiston vanhassa juhlasalissa S212 toukokuun 15. päivänä 2020 kello 12.

Academic dissertation to be publicly discussed, by permission of
the Faculty of Humanities and Social Sciences of the University of Jyväskylä,
in the Old Festival Hall S212 on May 15, 2020 at 12 o'clock noon.

JYVÄSKYLÄN YLIOPISTO
UNIVERSITY OF JYVÄSKYLÄ

JYVÄSKYLÄ 2020

Editors
Michael Freeman
Department of Language and Communication Studies
Timo Hautala
Open Science Centre, University of Jyväskylä

Permanent link to this publication: http://urn.fi/URN:ISBN:978-951-39-8130-3

ISBN 978-951-39-8130-3 (PDF)
URN:ISBN:978-951-39-8130-3
ISSN 2489-9003

Devoted to my beautiful wife,
Maryam,
without whom would have been no end to this journey

“”

ABSTRACT

Amin Rasti Behbahani
Investigating the effect of digital game tasks, inducing different levels of involvement load, on the acquisition of vocabulary items
Jyväskylä: University of Jyväskylä, 2020, 198 p.
(JYU Dissertations, ISSN 2489-9003; 210)
ISBN 978-951-39-8130-3
Summary
Diss.

In this empirical study, the effectiveness of digital game tasks, inducing different levels of involvement load, on the acquisition of vocabulary items were studied both quantitatively and qualitatively. Participants were 30 randomly recruited Persian speakers (14 males, and 16 females, aged 13 – 15 years). The research design included pre-tests, treatments, and post-tests. After the pre-tests, participants were randomly assigned to three involvement load groups, A, B, and C, containing 10 participants each. Concurrent think-aloud data were collected from two randomly selected pairs in each group. The digital game tasks designed for group A induced the lowest, the group B, a moderate, and the group C, the highest levels of involvement load. All participants played a commercial adventure digital game, *Haunted Hotel: Death Sentence*, in pairs by reading and following a game guide. From the game guide, 20 target words comprising inanimate object names or lexical nouns, were selected. At 3 weeks after task completion, the participants performed delayed post-tests. The quantitative data analysis showed that although digital game tasks can be effective in the acquisition of the scopes, and dimensions of a word, productive knowledge of the target words was superior to receptive knowledge. Moreover, the group B participants, counter to theoretical expectations, showed the poorest performance. The qualitative data analysis showed that, in performing digital game-based tasks, task structure, context, and strategy selection can all affect vocabulary acquisition. Moreover, participants employed distinct learning approaches that demanded the use of both universal moves (*information search, negotiation, turn taking,* and *trial-and-error*) and exclusive strategies (group A used *input enhancement strategies,* group B, *inferencing* and *hypothesis testing strategies,* and group C, *memory search, feedback request, word association strategies*, and *planning*). Hence, prospective teachers should be made aware of the predictive power of involvement load hypothesis.

Keywords: vocabulary acquisition, digital games, levels of involvement load, think-aloud, vocabulary learning strategies, task

TIIVISTELMÄ

Amin Rasti Behbahani
Investigating the effect of digital game tasks, inducing different levels of involvement load, on the acquisition of vocabulary items
Jyväskylä: University of Jyväskylä, 2020, 198 p.
(JYU Dissertations, ISSN 2489-9003; 210)
ISBN 978-951-39-8130-3
Summary
Diss.

Tutkimuksessa tarkasteltiin kvantitatiivisin ja kvalitatiivisin keinoin digitaalisten pelien vaikutusta sanaston oppimiseen. Tutkimushenkilöinä oli 30 persiankielistä 13-15 -vuotiasta nuorta (M=14; N=16), jotka pelasivat englanninkielistä kaupallista seikkailupeliä *Haunted Hotel: Death Sentence* pelioppaan avulla. Ensiksi kaikille tutkimushenkilöille suoritettiin sanaston laajuutta mittaava esitesti. Esitestin jälkeen tutkimushenkilöt jaettiin kolmeen ryhmään, joiden peliopasta oli muokattu siten, että esitystapa edellytti tutkimushenkilöiltä eriasteista "paneutuneisuustaakkaa" (involvement load). Täten tehtävän taso vaihteli ryhmä A:n matalasta, B:n keskitasoon ja C korkeaan tasoon. Kussakin ryhmässä oli 10 henkilöä ja sanojen oppimista tarkasteltiin 20 pelissä ja pelioppaassa esiintyvän substantiivin avulla. Kolme viikkoa pelaamistilanteen jälkeen tutkittavat osallistuivat sanaston reseptiivista ja produktiivista osaamista mittaavaan jälkitestiin (delayed post-test). Lisäksi kerättiin ääneenajattelu -aineistoa kunkin ryhmän kahdelta satunnaisesti valitulta parilta. Tulosten kvantitatiivinen analyysi osoitti, että pelit voivat vaikuttaa sanaston oppimisen ja tiettyjen piirteiden omaksumiseen positiivisesti, mutta myös että produktiivinen tieto hallittiin paremmin kuin reseptiivinen. Tilastollinen analyysi osoitti myös, että – hieman odotuksenvastaisesti - ryhmä B:n osallistujat selviytyivät heikoimmin. Tulosten laadullinen tarkastelu toi esiin, että pelaamiskontekstissa sanaston oppimiseen vaikuttavat useat eri tekijät, mm. osallistujien käyttämien strategioiden valinta. Tutkimushenkilöillä oli käytössä eri tyyppisiä lähestymistapoja, joista osa esiintyi kaikkien ryhmien toiminnassa (*informaation haku (esim. oppaasta), neuvottelu parin kanssa, "vetovastuun" vuorottelu parin kanssa,* sekä *yrityksen ja erehdyksen strategia*) ja osa puolestaan vain joissain ryhmissä. Siten esimerkiksi ainostaan ryhmä A hyödynsi sanan *ääneenlukua,* ryhmä B *päättelyn* ja *hypoteesintestauksen* strategioita ja ryhmä C *muistista hakua*, *palautepyyntöä, sana-assosiointia*, ja *etenemissuunnitelmaa*). Tutkimuksen tulokset antavat viitteitä myös esimerkiksi siihen, miten digitaalisia pelejä voitaisiin hyödyntää kielenopetuksessa ja millaiset tehtävät näyttäisivät olevan hyödyllisiä sanaston oppimiselle.

Keywords: Sanaston oppiminen, Digitaaliset pelit, Paneutuneisuustaakan tasot, Ääneen ajattelu, Sanastonoppimisstrategiat, Tehtävä

Author Amin Rasti Behbahani
Department of Languages and Communication Studies
University of Jyväskylä, Finland
amin.rasti@jyu.fi
0000-0002-1507-6986

Supervisors Senior Lecturer/Professor Katja Mäntylä, Ph.D.
Department of Languages and Communication Studies
University of Jyväskylä, Finland/
School of Languages and Translation Studies
University of Turku, Finland

Professor Emerita Hannele Dufva, Ph.D.
Department of Languages and Communication Studies
University of Jyväskylä, Finland

Reviewers Adjunct Professor (Docent) Maarit Mutta, Ph.D.
Department of French Studies
University of Turku, Finland

Assistant Professor Majid Fatahipour, Ph.D.
Department of English
Islamic Azad University, Parand, Iran

Opponent Adjunct Professor (Docent) Maarit Mutta, Ph.D.
Department of French Studies
University of Turku, Finland

ACKNOWLEDGEMENTS

Every journey has an end, and mine is here. Although it was a challenging one, it has ended well. However, this happy ending was not achieved without help and support.

First of all, I would like to acknowledge my supervisor, Katja Mäntylä, for being with me during all these years, for her on-demand supports, her challenging and valuable comments, her suggestions for improvements, and all the time and energy that she invested in reading, commenting on, and helping me in structuring this thesis. I acknowledge that she was right in insisting on including qualitative analysis in the study. Although this looked difficult and terrifying in the beginning, it not only helped me to add a whole new dimension to my work but also helped me develop new research skills. Thanks a million, Katja.

Next, my profoundest thanks go to Hannele Dufva, my advisor, whose smiling face is etched on my brain. Her comments, suggestions, support, and discussions were very constructive and insightful. You will be in my mind forever, Hannele. Thank you.

I would also like to express my very great appreciation to Ari Huhta, Laura Stark, Anne Pitkänen-Huhta, and Mika Lähteenmäki for their support during my academic and emotional trials. Thanks, Ari, for your academic assistance; thanks, Laura for your emotional and scientific support; thank you, Anne for your care; and thank you, Mika for your whole-hearted support. Without you, I would not have completed my thesis so soon.

I am also especially grateful for the assistance given by the study participants. Thank you for being available and showing interest in my research.

Finally, I would like to offer special thanks to my family for their non-stop emotional support and tolerance. Thanks mom and dad for unconditionally backing me in every situation, no matter what, during my life. And thank you, Maryam, my beautiful and patient wife, for tolerating loneliness when I was not in the mood for anything else but getting on with my thesis. Thank you, too, for your statistical assistances and consultations. Without your advice and help, I would not have easily figured out what was going on in my quantitative data and what those weird numbers and formulas had to offer.

Thank you, everyone Paljon kiitoksia kaikille از همه شما سپاسگزارم

Jyväskylä, 20.01.2020
Amin Rasti Behbahani

FIGURES

TABLES

CONTENTS

1 INTRODUCTION

Lack of vocabulary has always been a major challenge for second or foreign language learners (Kang, 1995). Vocabulary refers to the collection of words, and their definitions, that every language learner acquires during his efforts to master the target language (Carter, 1998). Evidently, vocabulary acquisition[1] is a crucial part of learning English. According to Nation (2006), knowledge of at least 8,000 to 9,000 words is necessary if one tends to comprehend a written English text; moreover, one needs to know 6,000 to 7,000 English words in order to understand a spoken form of communication. It seems reasonable, therefore, to deduce that language learning is not feasible without the acquisition of an adequate vocabulary (Kang, 1995).

However, multidimensionality of knowing a word has made vocabulary acquisition challenging; moreover, because acquisition of word knowledge takes place incrementally, vocabulary acquisition may not occur in just one sitting (Ringbom, 1987; Nation, 2001; Schmitt, 2007). Nation (2001) observes that knowing a word has two dimensions: when someone understands a word during reading or listening, which is called receptive knowledge, and when he uses a word in his writing or speaking, which is called productive knowledge. Nation (ibid.) also stipulates other aspects of knowing a word such as form, meaning, and use. Therefore, to overcome the complexity and challenges of vocabulary acquisition, the learner must know many interrelated bits of information.

This in turn requires extensive instruction. To date, a fully comprehensive vocabulary teaching/acquisition method has not been introduced. Instead, researchers and language teachers have put their trust in a wide range of methods, techniques, theories, and tools to assist language learners in accelerating their vocabulary acquisition. Recently, researchers' attention has been drawn to the potential of digital games in boosting both foreign and second language vocabulary acquisition.

---

[1] It is important to note the differences between acquisition and learning (Krashen, 2009). In this thesis, I discuss vocabulary acquisition rather than vocabulary learning. Therefore, the term vocabulary learning is not used unless the nature of vocabulary uptake has been reported as vocabulary learning in a referenced study.

Video or digital games have been shown to be beneficial for both language learning and acquisition. Studies on the potential and effectiveness of digital games in language learning have recently led some language learning researchers to see digital games as offering new opportunities and providing a safe virtual environment for experimentation with language learning (Kirriemuir, 2002). In the field of vocabulary acquisition, many studies have found digital games to be effective in vocabulary acquisition (Bakar & Nosratirad, 2013; Jasso 2012; Rankin, Gold, & Gooch, 2006b; Vahadat & Rasti Behbahani, 2013; Hung, 2011; Yip & Kwan 2006; Alias, Rosman, Rahman, & Dewitt, 2015, Chen, Tseng, Hsiao, 2018). Overall, these studies seem to consent in one point that general/field-specific vocabulary acquisition in educational/commercial digital game-based learning (DGBL) contexts can be often more extensive than previously applied vocabulary instruction methods (Fotouhi-Ghazvini, Earnshaw, Robinson, & Excell, 2009).

Although the digital game-based vocabulary acquisition literature has shed an optimistic light on the educational implication of digital games and has supported them as trustworthy utilities, especially for their effectiveness in form-meaning acquisition, the vocabulary acquisition process is multidimensional and includes dimensions and scopes, such as productive/receptive, and recall/recognition, respectively, that have been little researched. Moreover, recent findings on the effectiveness of digital games on vocabulary acquisition have also shown that it can be diminished and hampered by the internal elements of such games. deHaan, Reed, & Kuwada (2010) mention that interactivity in digital games is one of the constructive elements with the potential for improving language acquisition. However, they report that imbalance in the level of interactivity can have a negative effect on both vocabulary acquisition and word form recall. Although it has been found that more interactivity-rich digital games are potential candidates for better vocabulary acquisition (Zhonggen, 2018), deHaan et. al. (2010) note that uncontrolled high levels of interactivity in digital game tasks may lead to cognitive overload, or excess of mental processes over the limits of memory, hindering recall of vocabulary items. This finding is important because it warns that, if not controlled and monitored, even effective factors can have negative effects on vocabulary acquisition in digital game-based learning contexts.

Reynolds (2017) has studied the nature of the digital game-based vocabulary acquisition tasks closely. He reports that digital-game based vocabulary acquisition tasks appear to induce specific levels of involvement load by requiring actions such as need, search, and evaluation, which are constructive components of the involvement load hypothesis (Hulstijn & Laufer, 2001; see also section 2.5). In traditional vocabulary acquisition contexts, it has been reported that task-induced levels of involvement load can predict the success rate of every vocabulary acquisition task; moreover, higher levels of task-induced involvement load guarantee effective vocabulary acquisition (Hulstijn & Laufer, 2001; Kim, 2011). Thus, finding the most optimal task-induced level of involvement load in digital game-based vocabulary acquisition tasks would be

valuable as it increases the chance of language learners succeeding in increasing their vocabulary. However, despite his emphasis on controlling and monitoring the components of the involvement load hypothesis, Reynolds (2017) did not indicate the optimal level of involvement load for effective vocabulary acquisition in digital game tasks.

Given the importance of different levels of involvement load and the findings by Reynolds (2017) and deHaan et. al. (2010), it can be concluded that studying, monitoring, and controlling task-induced levels of involvement load are necessary for optimizing the effectiveness of digital game-based vocabulary acquisition tasks. Hence this study, which was designed to investigate the effect on the acquisition of target vocabulary items of digital game tasks inducing different levels of involvement load. Moreover, in this study, the investigation extends to the acquisition of form-meaning, and to the dimensions and scopes of the target vocabulary items, such as receptive (recall/recognition) and productive (recall/recognition) to consider the multidimensional nature of vocabulary acquisition, to contribute to filling the gap in the digital game-based vocabulary acquisition literature, and to taking the knowledge in the field one step further towards more effective vocabulary teaching and acquisition. The main aims of this study can be summarized as follow:

1. evaluating the effect of the digital game tasks on the acquisition of target vocabulary items, and knowledge of their dimensions.
2. Identifying the dimensions and scopes of word knowledge that are more effectively acquired after interaction with the digital game tasks.
3. Investigating the effect on the acquisition of target vocabulary items of digital game tasks inducing three different levels of involvement load.

To achieve these aims, through a quasi-experimental research method, and pre-test, treatment, post-test design, 30 volunteer Iranian participants were recruited. Pairing the participants provided the possibility of collecting qualitative data and applying concurrent think-aloud protocols. After the participants, in pairs, had performed the researcher-designed digital game-based vocabulary acquisition tasks for learning 20 English target words, which were the names of inanimate objects and lexical nouns, the effectiveness of the digital game task-induced involvement loads was measured, evaluated and monitored both quantitatively and qualitatively.

In chapters two and three I review the essentials of vocabulary acquisition, digital games, and previous studies that assist understanding the concepts used in this thesis. I discuss the importance of vocabulary acquisition and how digital games have evolved into beneficial tools for vocabulary acquisition. In the fourth chapter, i.e., methodology, I explain and describe the materials used in the study, including the digital game, namely *Haunted Hotel: Death Sentence – Collector's Edition*, and the game guides, the measurement instruments, such as the achievement tests, vocabulary size test and interview, and the participant demographics. I also describe and explain how the empirical study was conducted, including how the participants were categorized into three groups and the concurrent think-aloud data were collected and analyzed. In chapter five,

I present and discuss the results of both the qualitative and quantitative data analyses. In chapter six, I discuss the contribution of this study to the literature and what it adds to our knowledge of vocabulary acquisition.

2 VOCABULARY, ACQUISITION, TEACHING, AND CHALLENGES

Knowledge of vocabulary is of great importance for language acquisition and language learners. It is argued that

> Excellent reasons exist for devoting attention to vocabulary and spelling. First there are practical reasons. A large vocabulary is, of course, essential for mastery of a language. Second, language acquirers know this; they carry dictionaries with them, not grammar books, and regularly report that lack of vocabulary is a major problem. On the other hand, All other things being equal, learners with big vocabularies are more proficient in a wide range of language skills than learners with smaller vocabularies, and there is some evidence to support the view that vocabulary skills make a significant contribution to almost all aspects of L2 proficiency. (Krashen, 1989, p. 440)

Thus, vocabulary acquisition and teaching can be considered one of the major activities in every second or foreign language classroom. However, vocabulary acquisition and teaching are challenging. There are many factors that contribute to making vocabulary acquisition and teaching a challenging task. Thus, owing to the multidimensional nature of vocabulary and its challenges for learners, the researcher seeking to learn more about vocabulary acquisition and teaching must consider both theoretical and practical findings. Hence, I have tried to cover the major areas of research, practice and theory essential for conducting vocabulary acquisition research.

2.1 Learning vs. Acquisition

At the outset, it should be noted that in applied linguistics the concepts of learning and acquisition are defined as two distinct processes. Krashen (2009, p. 10) defines acquisition as "a subconscious process [,in which] language acquirers are not usually aware of the fact that they are acquiring language, but are only aware of the fact that they are using the language for communication", whereas

learning "refers to conscious knowledge of a second language, knowing the rules, being aware of them, and being able to talk about them. […] learning is "knowing about" a language, known to most people as "grammar", or "rules". " Despite the conceptual differences between these two terms, I have used them interchangeably throughout the thesis for two reasons: 1) the authors of the studies referred in the literature review describe their studies as either learning or acquisition studies; and 2) There are terms that are commonly known and introduced as either learning or acquisition, for example, *digital game-based learning*, *task based learning*, etc.

2.2 What is a word?

The first research challenge concerns the definition of a word. Language is made up of words. Words sit together to create phrases, sentences, and larger units of language. Singleton (1999) states that words are the main part of a language because they play a distinctive and crucial role in linguistic communication. But what precisely *is* a word?

Bloomfield (1933), defines a word as "a form which consists entirely of two or more lesser free forms… a free form which is not a phrase is a word… in brief, a word is a *minimum free form*" (Language, p. 178). He explained that a free form, unlike a bound form, is not a part of a larger form and can occur independently. Fries and Traver (1940, p. 87) state that "a word is a combination of sounds acting as a stimulus to bring into attention the experience to which it has become attached by use".

Carter (1998) defines a word by considering the different features of a word. According to him, the most commonsensical definition of a word would seem to be the orthographic definition. That is, a word is a combination of letters bounded by a space or punctuation mark on either side. However, he observes that this definition is not adequate; for example, if words like *sit*, *sat*, *sits*, and *sitting* are considered separate words, should they be separate entries in dictionaries? Moreover, there are words that have the same orthographic form but different meanings. Should we consider *bank*, as a financial firm, and *bank*, as a place near a river, as one word or two? Singleton (1999) also rejects the orthographic definition because he believes that this definition is only applicable to languages with a Roman or Cyrillic alphabet but not for languages like Chinese that is a tonic language and has a different alphabet. Accordingly, Carter (1998) speculates that

> An orthographic definition is one which is formalistic in the sense of being bound to the form of a word in a particular medium. It is not sensitive to distinctions of meaning or grammatical function. To this extent it is not complete. (Carter, 1998, p. 5)

Carter (1998, p.5) considers the definition of a word as "the minimum meaningful unit of language". Forms like *bank* are more acceptable as words in this definition

because "this definition presupposes clear relations between single words and the notion of meaning" (Carter, 1998, p.5). But what about forms like *bus station*, which contains two forms and one meaning? In addition, what about forms like s*hould, if,* and *could*, which can have different meanings in different contexts? To overcome this problem, Carter provides another definition for a word: a word will not have more than one stressed syllable. However, there are forms that not only do not convey meaning by themselves but also do not receive stress except in a specific situation (e.g. *them*, *but, by*).

Thus, Carter (1998) concludes that defining a word is very problematic as none of the definitions, either commonsensical or technical, gets us far. He summarizes the problems in defining a word as follows: 1) there are words that do not fit into the orthographic, free form, or stressed-based definition of a word; 2) considering words as units of meaning is vague and asymmetrical; 3) different forms of a word do not count as different words; 4) there are words that have the same forms but completely unrelated and different meanings; 5) the existence of idioms further complicates any attempt at defining a word. Singleton (1999) and Milton (2009), in turn, state that although words are a vital part of a language, providing a comprehensive definition of a word is very challenging.

Despite the controversy and challenges presented by defining a word, I think a working definition can be formulated, drawing on the definitions presented above, by considering the context in which the word is used. In my study, I have selected words that are linguistically called nouns. In addition, they refer to inanimate objects. Thus, generally, in the context of my study, a word is a combination of sounds and syllables that has a pre-identified orthographic boundary. I also recognize the boundary as "the minimum meaningful unit of language" if it forms a noun. Thus, for the purposes of this study, I define a word as a combination of sounds and syllables in a specific orthographical form that labels an inanimate object and can grammatically be categorized as a noun.

2.3 What is it to know a word?

An important question, in second/foreign language vocabulary acquisition studies, is what is understood as knowing a word. Given the challenges concerning the definition of a word, it is also hard to define what knowing a word is, although valuable attempts have been made. If you ask what it means to know a word, the average educated person may answer that it means knowing the spelling and meaning of the word (Schmitt, 2010b). In general, learners think that knowing a word means knowing its correct spelling, pronunciation and meaning (Nation, 1990, 2001). These answers can be considered reasonable. In other words, knowledge of the written/spoken form and meaning of a word is the basic form of word knowledge (Schmitt, 2010b).

Attempts from the applied linguistics point of view by Richards (1976), Ringbom (1987) and Nation (2001) at defining what knowing a word means show that this question is more challenging and demands a more profound answer.

According to Henriksen (1999, as cited in Nation & Webb, 2010), every individual has a lexical competence that comprises three dimensions: 1) partial to precise knowledge of a word, 2) how profoundly the individual knows the word, and 3) being able to use the word in both speaking/writing and listening/reading. In other words, knowing a word means knowing it in each of the three aspects of lexical competence. The more competent one's lexical knowledge, the more profoundly one knows a word. However, Henriksen appears to be merely scratching the surface, while other scholars have dug deeper into what knowing a word in a second/foreign language means. For instance, Ringbom (1987) sees L2 lexical knowledge as a complex interconnected matrix of knowledge systems that are accessed for both comprehension and production. He posits that when an L2 language learner wants to learn a word, he is faced with different linguistic tasks such as learning the internal form (morphology), the meaning (semantics), the use of the word in a sentence (syntax), the words that it can be combined with (collocation), the words that are related to it (association), and, finally, the extent to which the word can be accessed (accessibility). Ringbom (1987) describes knowing a word as a continuum from no knowledge in the early stages of learning through incremental increases in knowledge to knowledge at an advanced level. At the advanced level of word knowledge, a learner has complex L2 lexical knowledge and has stored a lot of information about each word in his lexicon. Thus, it can be deduced that knowing a word is challenging because it involves dealing with a lot of information. Ringbom's description of the continuum and of the complex nature of knowing a word is presented in the following table:

TABLE 1 Complex nature of knowing a word (Adapted from Ringbom, 1987, p. 37)

Beginner Level	*Incremental Development* →		Advanced Level
accessibility	the word is accessible within specific context only		the word is accessed regardless of context
morphology	knows one form of word	knows words in all its forms (spoken, written, inflected)	knows the possible derivations of a word
syntax	knows no syntactic constraints	knows some constraints	knows all syntactic constraints
semantics	knows approximate meaning only	knows one meaning only	knows all possible meanings
collocation	knows no collocational constraints	knows some constraints	knows all collocational constraints
association	knows no associative constraints	Knows some constraints	knows all associative constraints

Richards (1976) in turn enlists different aspects of knowing a word in second language acquisition. He regards form-meaning knowledge as the basic and first step in knowing a word, as this knowledge alone does not help in using the word appropriately and confidently in a range of different contexts. In other words, L2 word knowledge should be considered as a range of knowledges Richards (1976), in his seminal paper, explains knowing a word by positing eight assumptions:

- the process of knowing a word does not stop for a human by maturing and getting older (first assumption)
- knowing a word means knowing the frequency and degree of exposure to a word and its associations (second assumption)
- knowing a word means knowing its functions and the situations in which it is used (third assumption)
- knowing a word means knowing how that word behaves syntactically (fourth assumption)
- knowing a word involves knowledge of its form and derivations (fifth assumption)
- knowing a word entails knowledge of its associations and how it is associated with other words (sixth assumption)
- knowing a word is knowing semantic aspects of a word (seventh assumption)
- knowing a word means knowing many different meanings of that word (eighth assumption)

Although Richards' assumptions are valuable, they are not systematic. Therefore, based on Richards' (1976) assumptions, Nation (1990, as cited in Schmitt, 2010b) proposed a concise, refined, and systematic version of L2 word knowledge. In Nation's first attempt, knowing a word means knowing its meaning, written form, spoken form, grammatical characteristics, collocations, register constraints, frequency and associations (Nation, 1990, p. 31). In a newer version, Nation (2001) provides a convincing and systematic definition of knowing a word. He also considers the active/passive dimension, which he renames as receptive/productive. Nation sees this distinction as applicable to different kinds of language knowledge and use. Thus, when applied to vocabulary, they cover all the aspects of what is involved in knowing a word (2001, p. 26). On the notion of active/passive, he defines receptive knowledge of a word as recognizing and understanding the word when it is read or heard. Productive knowledge of a word includes not only receptive knowledge but also knowledge of spelling, pronunciation, grammatical usage, functional use, collocations and synonyms (Nation, 1990, 2001). Finally, he adds that "At the most general level, knowing a word involves form, meaning, and use" (Nation, 2001, p. 26). Nation provides the following table to explain different aspects and dimensions of what knowing a word comprehensively involves.

TABLE 2 Aspects Involved in Knowing a Word (adapted from Nation, 2001, p. 27)

Form	Spoken	R	What does the word sound like?
		P	How is the word pronounced?
	Written	R	What does the word look like?
		P	How is the word written and spelled?
	Word parts	R	What parts are recognizable in this word?
		P	What word parts are needed to express the meanings?
Meaning	Form and meaning	R	What meaning does this word form signal?
		P	What word form can be used to express this meaning?
	Concept and referents	R	What is included in the concept?
		P	What items can the concept refer to?
	Associations	R	What other words does this make us think of?
		P	What other words could we use instead of this one?
Use	Grammatical functions	R	In what patterns does the word occur?
		P	In what pattern must we use this word?
	Collocations	R	What words or type of words occur with this one?
		P	What words or type of words must we use with this?
	Constraints on use (register, frequency, etc.)	R	Where, when, and how often would we expect to meet this word?
		P	Where, when, and how often can we use this word?
	R = Receptive Knowledge P = Productive Knowledge.		

The receptive/productive dimension has been considered to encompass two scopes, namely, recognition and recall, especially, in the form-meaning aspect of word knowledge (Nation, 2001; Laufer & Goldstein, 2004). Accordingly, the form-meaning aspect can also include knowledge of productive recall (retrieving L2 word forms by their L1 definitions), productive recognition (recognizing dictated L2 words), receptive recall (retrieving definitions of L2 words by their forms), and receptive recognition (recognizing the most relevant definitions of L2 words among many other definitions).

TABLE 3 Scopes of form-meaning knowledge of vocabulary items (adapted from Laufer & Goldstein, 2004, p. 407)

	Recall	Recognition
Productive (retrieval of form)	Supply the L2 word	Select the L2 word
Receptive (retrieval of meaning)	Supply the L1 word	Select the L1 word

The acquisition of aspects, dimensions and scopes of word knowledge have also been previously studied and discussed. Although the precedence of receptive

over productive knowledge in the acquisition of dimensions of word knowledge has been discussed (Morton, 1977 as cited in Barcroft, 2004; Meara, 1997; Nation, 1990, 2001; Schmitt, 2008, 2010a), vocabulary acquisition studies showing that the acquisition of productive word knowledge precedes that of receptive word knowledge have also been published. For example, de la Fuente (2002), investigated the effect of three different tasks on the acquisition of target words. The tasks, which she named conditions, were "non-negotiated premodified input (NNPI), negotiation without pushed output (NIWO), and negotiation plus pushed output (NIPO)" (p. 81). Her participants were 32 L2-learner volunteers who were studying Spanish in Georgetown University. She randomly assigned them to three groups based on the three conditions. She selected 10 Spanish target words that participants had not previously been exposed to. During two sessions, participants performed two listening tasks in which they were to listen to target words and locate relevant objects or pictures in the room. The NNPI participants had no rights to ask any questions, The NIWO participants could negotiate the meaning of the target words with their native speaker partners (NSs) for 1 minute only. Finally, the NIPO participants were to provide information for the NSs to find the objects or pictures. The NIPO participants were also allowed to ask the NSs questions. This task was repeated in the same manner a day later but with the inclusion of time-on-task. Three post-tests, which measured both receptive and productive knowledge of the target words, were administered three times as both immediately after and at 3 weeks thereafter. Comparison of the results revealed that task type of task was a defining factor in the acquisition of dimensions of word knowledge. It was only in the negotiated interaction plus output (NIPO) condition that the participants significantly acquired both receptive and productive knowledge of the target words, although, surprisingly, productive acquisition preceded receptive acquisition of the target words. Webb (2007a) studied the effect of contextualized and decontextualized vocabulary learning tasks on the acquisition of different aspects and dimensions of word knowledge. He recruited 84 Japanese EFL students who has scored 80% in the version 1 Vocabulary Level Test, which measures receptive knowledge of the first 2000 most frequent words. He randomly assigned them into experimental and comparison groups. The experimental group were administered the target words in glossed sentences while the comparison group had them in word pairs. A surprise test, which measured knowledge of orthography, pragmatic association, meaning and form, syntagmatic association, and grammatical function was administered after they had completed their primary tasks. Although comparison of the results showed no significant difference in gains in aspects of word knowledge between the two groups, he reported that the largest gains were found for the productive knowledge of meaning. Mondria and Wiersma (2004, pp. 85-86) discuss factors such as "overlap between receptive learning and productive learning" and "the decay of receptive knowledge vs. productive knowledge" as the reasons for such inconsistencies in the literature. For example, they argue that although a certain amount of productive knowledge can be gained from doing receptive learning tasks and vice versa, "the receptive

retention as a result of productive learning in general lags behind the receptive retention as a result of receptive learning"(p. 85). Based on this finding, it can be understood why, in the study by de la Fuente (2002), the NIPO participants outperformed the others in the productive tests. Moreover, Mondria and Wiersma (2004) reported that the rate of retention loss in receptive knowledge has been found to be higher than that in productive knowledge. Therefore, the participants in the study by Webb (2007a) might have experienced more attrition in their receptive than productive knowledge. Hence, simultaneous receptive and productive vocabulary acquisition is recommended because it elevates the acquisition of these aspects of word knowledge (Gass, 1999; Lee & Muncie, 2006) by diminishing the chance of form processing before meaning processing, which has been reported to be a negative factor in the acquisition of target words and aspects of these (Ellis & He, 1999).

Overall, knowing a word in second/foreign language acquisition includes knowledge of lots of bits of information that are systematically interrelated. It is more like a continuum than the known versus unknown dichotomy or mere form-meaning link. This knowledge ranges from zero through partial to precise knowledge and is true for all aspects of knowing a word (Schmitt, 2010b; Ringbom, 1987). Therefore, when one knows a word, one knows most of the aspects, dimensions, and scopes of knowing a word listed in the above table. In general, in this study, I consider a word as known if I can find even the slightest development in the form–meaning relationship, which constitutes the basic knowledge of any type of word.

2.4 What is it to acquire a word?

Different answers from different perspectives have been given to the question of what acquiring a word is. Some psychological studies have posited that "in essence, the process of learning a foreign language word is to map a novel sound pattern to a particular semantic field that may have an exact equivalent in the native language" (Ellis & Beaton, 1993a, p. 560). In turn, phonological, studies on cognition and memory often consider new vocabulary acquisition as achieved through gaining knowledge of how the available sounds in the target language are linked together to represent a particular instance or specify a class of instances (Gathercole & Baddeley, 1990). Moreover, these studies have placed more emphasis on phonological memory and hold that acquiring vocabulary means learning how to sequence phonological properties, such as phonotactic sequences, syllable structures, and the categorical units of a language (Ellis & Sinclair, 1996).

However, with regard to the challenging points discussed above about defining a word and knowing a word, applied linguists believe that vocabulary acquisition is an incremental process in which one cannot learn a word in a single exposure (Schmitt, 2007, 2008, 2010). Taking Nation's (2001) definition of knowing a word, they report that some of these features are acquired before other

features and that knowledge of some aspects does not guarantee knowledge of other aspects (Schmitt, 2007). Thus, in applied linguistics, vocabulary acquisition is defined as a gradual process during which different aspects of a word are acquired. Applied linguists also conceive vocabulary acquisition as a continuum rather than a known/unknown dichotomy. For example, Schmitt (2007), as discussed by Ringbom (1987) earlier, illustrates learning a word in a second language as an arrow that proceeds from no knowledge to full mastery and proposes that this continuum applies to every aspect of acquiring a word. He exemplifies the incremental learning of the written form of a word as follows:

FIGURE 1 Incremental Nature of Learning a Word (adapted from Schmitt, 2007, p. 749)

Ellis (1995) summarizes the process by stating that to acquire a foreign language word we must first recognize it as a word. Thus, there must be different lexicons (or knowledge bases) in our mind with different channels of input and output. He suggests that to understand speech, the auditory input lexicon must be able to categorize sounds in sequences that sound meaningful for us; to understand a word when we read it, the visual input lexicon must be able to recognize orthographic patterns; to say a word, the pronunciation output lexicon must activate the articulatory organs to deliver a pre-sequenced pattern of sounds to be pronounced correctly; and that to write a word, the spelling output lexicon must provide a logical orthographic pattern. We must also learn the word's syntactic structure, semantic properties, its relation to other words, its place in the lexical order, and its referential properties. Thus, learning a word is a complicated but gradual process that involves many different logical, psychological and pedagogical processes (Schmitt, 2007; & Ellis, 1995). In a vocabulary acquisition study, the researcher must consider this complexity so that he can either control or enforce the effect of the multiple factors that play determining roles in vocabulary acquisition process.

2.4.1 Psycholinguistics and acquiring a word

Despite the extensive literature on learning a word, no concise theory exists for modelling the learning process itself (Nation, 1990, 2001; Nation & Webb, 2010; Schmitt, 2010a). However, many attempts have been made in psycholinguistics to show how a word is recognized, processed, stored, and accessed. In this subsection, I discuss two proposed models from the word acquisition point of view. I have selected these two models on the grounds of the important role they assign the lexicon and its internal processes during language use.

2.4.1.1 Levelt's model of lexical access in speech production

Levelt's model of lexical access in speech production, while in the first place a model of L1 language production, is of great interest for L2 vocabulary scholars owing to its heavy reliance on vocabulary knowledge and its deep involvement during language use. According to this model, the lexicon contains knowledge that is declarative in nature. In other words, a collection of facts and information are stored in individuals' memories. They can build or rebuild their collection by formal study or incidental learning. Moreover, the grammar and phonology of sentences are determined by the selection of words from the lexicon (Levelt, 1992). By the same token, it can be inferred that if the words in the lexicon determine the syntax used, other aspects of knowing a word might also play crucial roles in language production. Moreover, this view emphasizes that exposure to words in use is an influential way of developing vocabulary knowledge. Furthermore, it proposes that the decontextualized learning of vocabulary cannot be very effective even if it may be of some use in acquiring a word (Kang, 1995; Nation, 2001).

Levelt (1989, as cited in Nation, 2001) argues that, the lexicon contains two knowledge components in which forms and lemmas are stored separately for all types of words. Levelt, Roelofs, and Mayer (1999, p. 37) explain that the term lemma, first introduced by Kempen and Hoenkamp (1987) and later adopted by Levelt, was used to "denote the word as a semantic/syntactic entity" as opposed to the term lexeme that "denotes the word's phonological features". However, in his theory, Levelt later limited the concept to syntactic knowledge. Every lemma contains many bits of information such as semantic and grammatical knowledge, "that is, knowledge of the meaning components of a word and knowledge of the syntactic category (part of speech) of a word" (Nation, 2001, p. 38), syntactic category, grammatical functions, grammatical restrictions (ibid.) (Figure 2). Levelt (1992) adds that pointers link the morpho-phonological form of the word to the information contained in the lemmas; or, put more simply, written/spoken forms and meanings are linked in the lexicon by the pointers.

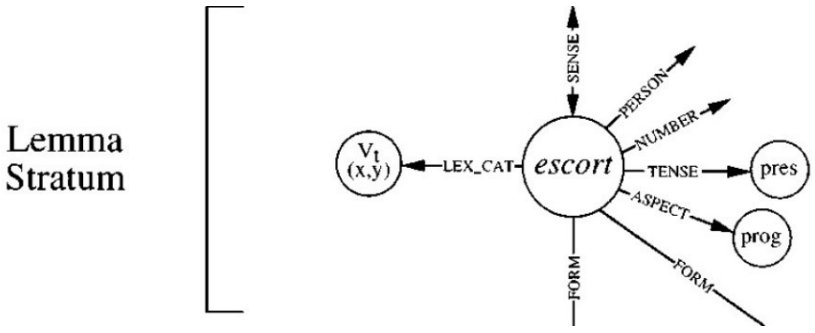


FIGURE 2 Information about the word *escort* in Lemma (adapted from Levelt, Roelofs, & Mayer, 1999, p. 4)

Levelt (1992) also states that there are not only internal relations in each entry but also there are external relations among entries. The internal relations of an entry's bits of information are essential owing to the cause-effect nature of language

reception and production. This type of relation means the existence of derivative forms and word families. On the external relationships between entries, Levelt (1992) distinguishes two type of relations: intrinsic and associative. In the intrinsic external relationship, entries are related based on four features: meaning, grammar, morphology and phonology. Semantic relations like synonymy, antonymy etc. derive from this type of relationship between entries. Associative relations are the main causes of collocations.

I see Levelt's model as highly relevant to the nature of vocabulary acquisition since, as discussed in the previous section, knowing a word is the outcome of the interaction of various factors and hence a complex incremental process. Levelt's model suggests that knowing and acquiring a word is highly important in language learning because of the role played in all language use by the lexicon. Moreover, the model supports the idea of the incremental nature of learning words. In this model, the lexicon is described as a world of bits of information that are related, linked and connected to each other in various and specific ways. Thus, to build such a lexicon for another language is time-consuming, and many processes, such as the building, rebuilding, and refining of both entries and links, must be invoked a great number of times.

2.4.1.2 Meara's model of word learning

In his model of word learning, Meara assumes that acquisition of a word

> consists of the building of a connection between a newly encountered word and a word that already exists in the learner's lexicon. This connection might be a link between the new L2 word and its L1 translation, or it might be a link between the new L2 word and an already known L2 word. (Meara, 1997, p. 118)

He also presumes the link is unidirectional; that is, he presumes that only newly acquired L2 words can activate or retrieve their specific L1 translations or L2 synonyms in the learner's lexicon while L1 translations or L2 synonyms may not be able to activate or retrieve the newly acquired L2 words. In this model, as Meara claims, vocabulary acquisition is considered a cumulative activity. He defines unknown words as words that have no connection to the learner's existing lexicon and known words as words that have connections to the learner's lexicon. However, they are different in the type and number of their connections. By the same token, the number of connections determines the strength of knowing a word. A word with great number of connections will be well known, but a word with a small number of connections will be poorly known. Meara adds that frequency of exposure has an important role in his model. More frequent exposure to a word increases and enriches the number of connections. He calls every exposure of a word, though small, an event and believes that, based on his model, these small events add up and, in the long run, build a large lexicon.

Although Meara states that "there are huge problems with this type of thinking, and I am not suggesting that the simple ideas for models that I have developed here should be taken really seriously" (1997, p. 120), Schmitt (2010a)

points out that this model could provide a fairly convincing description of the active/passive or productive/receptive states of words. Accordingly, Schmitt (2010a) explains that an item in the lexicon is active when it is connected to a productive item. This type of connection "lights up" the item for productive use. But for the receptively known items, there are no "incoming links from the lexicon" and they must be activated by an external stimulus. When activated, they can be recalled.

Schmitt (2010a) concludes that, according to this model, "the move from receptive to productive mastery is the results of a fundamental change in the way a lexical item is integrated into the mental lexicon" (p. 81). Schmitt (2010a) believes that this model can explain how it is possible for some words to be acquired productively in the short run despite only little input. It can also explain why some words are known productively for some time only and not later. According to this model, the reason is that they are no longer connected to productive items in the lexicon and thus no longer productively accessible.

Although these models do not cover the learning of a word comprehensively, they demonstrate that learning a word is a very complicated process that depends on many factors.

2.4.2 Processes and modes of acquiring a word

What happens when one acquires a word? What factors are involved during this mental process? In this section, I summarize the research on the processes underlying the acquisition of a word.

2.4.2.1 Processes involved in learning a word

Nation (2001) introduces three major processes, at least one of which should be beneficial and effective in any word acquisition situation. Although the presence of all three processes simultaneously is not essential, word learning activities should be designed in such ways that these three psychological conditions have the potential to assist learners in reaching their goal, i.e., acquiring a word. The three processes through which word acquisition can be started, practiced and guaranteed are n*oticing*, *retrieving*, and *generation*. Although Nation (2001) claims that noticing is the least effective and generation the most effective one of these, noticing is considered the preliminary stage from which learning starts (Truscott, 1998; Cross, 2002). Moreover, noticing both the form and meaning of the target word is a necessity for generating form-meaning links in vocabulary tasks (Ellis, 1994). Noticing refers to "seeing the word as an item to be learned" (Nation, 2001, p. 221). Through noticing, a learner gains awareness of the importance of the item as a useful language unit. Two important enabling factors for noticing are motivation and interest (Nation, 2001). Research also seems to support the effect of noticing on effective vocabulary acquisition. For example, Alcon (2007) found that techniques like a pre-emptive focus on form episodes that elevate the level of noticing are more effective in vocabulary learning than techniques that do not

elevate the level of noticing. In my study, the role of noticing has been enhanced by bolding the target vocabulary items.

Retrieving refers to both recognizing and recalling previously encountered words. Retrieval strengthens the connection between the cue and word knowledge. With respect to the receptive/productive aspects of acquiring a word, retrieval could be "receptive/productive, oral/visual, overt/covert, and in context/ decontextualized" (Nation, 2001, p. 221). Receptive retrieval refers to remembering the meaning when the written or spoken form is encountered. Productive retrieval is remembering the written or spoken form when language is used for communication. Two major factors that affect the retrieval of a word: "the learner's vocabulary size, and the length of time that the memory of a meeting with a word lasts" (Nation, 2001, p. 67). The results of word-repetition studies support the expected effect of retrieval on word learning. Webb (2007b), in a carefully controlled design, studied the effect of 1, 3, 7 and 10 encounters on vocabulary learning among 121 Japanese language learners. He measured the knowledge of form, meaning, orthography, grammatical functions, association, and syntax for the target words with 10 tests, controlling for the type and aspects of contexts and the participants' language level and proficiency. He found that more repetition (retrieval in any form for any aspect of knowing a word) resulted in better and deeper knowledge of a word.

The process considered the most effective of the three is generation. Generative use of a word refers to the use of a previously encountered word in another context, another derivation, or a in a way that is different from previously exposures. "Generative use is not restricted to metaphorical extension of word meaning and can apply to a range of variations from inflection through collocation and grammatical context to reference and meaning" (Nation, 2001, p. 69). Generation can also be receptive and productive. Receptive generation refers to encountering a word in a distinctive way, different from previous exposure in listening or reading. Productive generation refers to a new way of using a previously encountered word that best fits the context. Generation can also be understood as a matter of degree. Generation is low if the language context shows little change; for example, if *old friend* is replaced by *very old friend*. Generation is high if the word which has been met before is used in a completely new way; For example, if *old friend* is replaced by *my very old man* or *my very old boy*. To test this process, Joe (1998) studied the effect of text-based tasks and background knowledge (the ability to use a new word generatively and vocabulary knowledge) on incidental vocabulary acquisition. Her results showed that a greater level of generative use led to a greater gain in vocabulary items.

In this study, the importance and effect of those three processes, i.e. noticing, retrieval, and generation, are considered, controlled for, and tested. For example, in the task design, to be described later, retrieval and generation as well as noticing processes play a crucial role in assisting participants to acquire the target vocabulary items.

2.4.2.2 Implicit and explicit learning

Psycholinguists frequently assume that "humans are endowed with two dissociable learning abilities: *implicit* and *explicit* learning" (Ellis, 1995, p. 5). The assumption of these two abilities derives from connectionism (Laufer & Hulstijn, 2001). The connectionist definition of implicit and explicit learning is as follows:

> Implicit learning is acquisition of knowledge about the underlying structure of a complex stimulus environment by a process which takes place naturally, simply and without conscious operations. Explicit learning is a more conscious operation where the individual makes and tests hypotheses in a search for structure. (Ellis, 1995, p. 6)

Many applied linguists believe that such abilities are involved in learning a second or foreign language. Ellis (2009, p. 7) assumes that in SLA, "explicit language learning is necessarily a conscious process and is generally intentional as well. It is conscious learning where the individual makes and tests hypotheses in a search for structure […] it is a conscious, deliberative process of concept formation and concept linking". However, "implicit language learning takes place […] without any metalinguistic awareness. That is, the processes responsible for the integration of material into learner's interlanguage system and the restructuring this might entail take place autonomously and without conscious control" (p. 7).

Ellis's definitions of these terms, while broad, are applicable in second language acquisition. However, in the case of vocabulary acquisition and knowledge, psycholinguists limit the scope and redefine the concepts of implicit and explicit learning. They describe the implicit learning of a word as the acquisition of different meanings of a word by several exposures to that word in different contexts in cases when this happens completely unconsciously. In turn, they delineate the explicit learning of a word as a completely conscious process in which metacognitive strategies are used to facilitate its acquisition (Ellis, 1995).

2.4.2.3 Incidental and intentional vocabulary learning

In vocabulary acquisition research, the terms incidental and intentional acquisition are used instead of implicit and explicit learning. In other words, a word is acquired either incidentally or intentionally. Laufer and Hulstijn (2001) and Hulstijn (2001) observe that these two terms, incidental and intentional acquisition, have their origin in experimental psychology research in the beginning of twentieth century when they were used in explaining processes that differed from explicit and implicit learning.

From a pedagogical perspective, Schmidt (1994, as cited in Laufer & Hulstijn, 2001) refers to incidental vocabulary acquisition as the unintentional acquisition of vocabulary items in a situation where the primary aim is to learn, for instance, how to communicate. Hulstijn (2001) distinguishes these two terms operationally. The difference between incidental and intentional as can be seen by considering the use of a pre-learning instruction that warns learners about a retention vocabulary test to be administered later. Although Hulstijn (2001) states that while providing a precise distinctive definition of these two terms is not an

easy task, the incidental acquisition of a word refers to when a word is acquired while the learner is engaged in a language activity in which the main aim is not acquiring that word. In contrast, every activity that is specifically designed to add a word to the mental lexicon can be considered intentional word learning. Other researchers' definitions of incidental and intentional word learning mostly converge with that of Schmidt and Hulstijn (Laufer & Hulstijn, 2001; & Ahmad, 2011). Ahmad (2001), however, sees the intentional acquisition of a word as another case of rote learning since, during this process, the role of context is disregarded.

It is easy to agree with Hulstijn that it is not easy to define incidental and intentional vocabulary acquisition. In fact, applied linguists have proposed these two distinctive terms because they believe that, despite their overlap in meaning, the two notions should not be confused with implicit and explicit learning in memory studies (Laufer & Hulstijn, 2001). For example, while implicit learning can only be incidental, explicit learning can occur both incidentally and intentionally. Putting the various definitions provided by psychologists, psycholinguists, and applied linguists together, it can be stated that "Attention and deep focus on form-meaning connection are factors that can determine the type of learning to be incidental or not" (Huckin & Coady, 1999, p. 183).

2.4.2.4 The effectiveness of Incidental or intentional word learning

But which mode of word learning is more effective? Ahmad (2001) found incidental word learning from context to be an effective way of learning vocabulary, a motivator for extensive reading, and a promoter of deeper mental processing. In his study, incidental word learners also outperformed intentional word learners.

For the incidental acquisition of a word to happen, several exposures to the word are needed (Huckin & Coady, 1999). For example, it is claimed that for a word to be acquired incidentally, 10 exposures are desirable (Saragi, Nation, & Meister, 1978, as cited in Huckin & Coady, 1999). Although this might be true for first language vocabulary acquisition (Jenkins, Stein, & Wysocki, 1984), it may not be as true for second language vocabulary acquisition. Gathercole and Baddeley (1990) studied the number of exposures in second language vocabulary acquisition and reported that an increase in the number of exposures was effective. Nation (2001), in turn, states that 7 exposures might be enough for acquiring new words. The idea of the optimum number of exposures was tested systematically in the second language learning context by Webb (2007), who found that "If learners encounter unknown words ten times in context, sizeable [incidental] learning gains may occur. However, to develop full knowledge of a word more than ten repetitions may be needed" (p. 46).

Huckin & Coady (1999), who presume that guessing is the major form of incidental word acquisition, posit that incidental word acquisition also has its limitations. In incidental vocabulary acquisition, guessing is imprecise: "accurate guessing requires accurate word recognition and careful monitoring because there are many deceptive lexical items that can easily mislead the learner" (p. 189).

Guessing also needs a lot of time, as word acquisition will not take place if the context is not well understood. Learners without a good knowledge of (reading) strategies cannot learn words incidentally and effectively. Thus, guessing does not mean incidental acquisition, and multiword lexical items are not learned effectively through either guessing or incidental word acquisition.

Schmitt (2010b) reviewed both modes of word learning and listed their benefits. He argues that intentional word learning is beneficial because it generally leads to more robust and faster learning, it generally engages the learner more profoundly - which helps better retention - and it can help the teacher to select the words to be taught. For incidental learning, he wrote that it could be helpful in the case of words that teacher could not teach intentionally, it might provide contextual information that teacher could not teach easily and explicitly, it might be good for retrieving a word taught explicitly, and it could help word learning while developing other language skills.

When considering the benefits and limitations of both incidental and intentional vocabulary acquisition, scholars commonly recommend that incidental and intentional modes of vocabulary acquisition are implemented simultaneously. They must also be treated as complementary activities (Schmitt, 2007; & Hulstijn, 2001). Many studies support this idea. Zimmerman (1994) found that sub-technical words were better acquired if the learner received interactive word instruction 3 hours a week compared to reading alone. Qian (1996) found that decontextualized L2 vocabulary acquisition plus feedback was more effective than contextualized L2 vocabulary acquisition. Furthermore, in a study by Paribakht and Wesche (1997), reading plus vocabulary instruction groups outperformed, both qualitatively and quantitatively, reading-only groups in the final retention test. Hulstijn (2001), however, argues that modes are useful only if they provide enough information. He believes the mode is used for teaching or learning a word is of little importance; what matters is how to increase the quality of information processing regarding a target word, for example, by enhancing the learner's sense of the relevance of the target words to the task (Laufer, 2001).

Likewise, in this study, the debate on the boundary between incidental and intentional vocabulary acquisition and the advantages of applying both modes informed the design of the present main task, in which the target words are shown in bold while the aim is not explicitly to teach them. This will be discussed later.

2.4.3 What techniques are used for teaching and acquiring vocabulary?

Many efforts at decrypting the nature of learning a word, and many contributions to developing theoretical understanding of word acquisition have been made. However, this information is not useful unless it assists language learners to acquire a target word in practice. At this juncture, a brief review of attempts at translating theory into practice in order to effectively teach acquire words is presented.

Sökmen (1997) lists the following techniques as current trends in teaching and acquiring second language vocabulary:

1. Inferring from the context: inferring from the context mainly happens when acquiring words incidentally. In this technique, learners are encouraged to guess the meaning using contextual clues and the surrounding words. Fraser (1999) found inferencing techniques to be an effective lexical processing strategy for enhancing vocabulary acquisition. In his think-aloud study, applying a repeated measures design, he gathered introspective data from 8 francophone intermediate ESL learners while they were reading 8 texts over 5 months to monitor their strategies when engaging with unfamiliar words. By a cued-recall task, he tested the participants' rate of vocabulary acquisition. He found that the participants utilized lexical processing strategies, such as consulting, inferencing, ignoring, and not paying attention, that had different effects on their rate of vocabulary acquisition. He found that, among the implemented lexical processing strategies, inferencing was a significantly effective strategy for incidental vocabulary acquisition. Although inferencing may be an effective technique, it had its limitations and needed to be practiced with intentional techniques (Huckin & Coady, 1999; & Schmitt, 2010b). According to Nassaji (2003, p. 648), three types of cognitive processes are involved in successful inferencing: "a generator, an evaluator, and a metalinguistic control component". He explains the roles of generator and evaluator processes as processes that generate hypotheses about the meaning of a word and evaluate them. He defines metalinguistic processes as a series of steps and decision-making processes that run between the generating and evaluating of hypotheses about the meaning of a word. In a think–aloud study, he searched for the inferencing techniques that are the outcomes of these processes among 21 intermediate ESL learners while reading a passage. He also considered the role of learners' knowledge sources. He found that the learners used the following inferencing techniques when they invoked the generator, elevator, and metalinguistic control components during reading: repeating, verifying, self-inquiry, analyzing, monitoring, and analogy. Moreover, he found that the effectiveness of these inferencing techniques can be low if learners use only "the strategies and knowledge sources they [have] in their disposal" (Nassaji, 2003, p. 645). Finally, he reported that the quality of the inferencing strategies applied is more important than their quantity. Nassaji's findings aside, it seems the effectiveness of inferencing techniques on incidental acquisition of vocabulary item also depends on other factors such as context. Texts, for example, provide richer contextual cues, and enhance the effectiveness of both vocabulary and grammar acquisition (Hulstijn, Hollander, & Greidanus, 1996; Romos & Dario, 2015). For instance, Webb (2008), found that the quality of the context is even more influential on the lexical encoding process than frequency of exposure. In other words, rich contexts enrich word processing by providing more cues about the various aspects of words (Rott & Williams, 2003). Thus, inferencing techniques can be more successful because the quality of inferencing is high in such contexts.

2. Teaching most frequent words: word frequency is one of the most important factors in teaching and acquiring a word. Learning the most frequent words is an effective way for learners to improve their skills. For example, Nation (2006)

found that for full comprehension of a text knowledge of 98% of the word used was needed. To reach this level of comprehension with written texts, a learner of English must know at least the first 8 000 most frequent English words and with spoken texts the first 6 000 most frequent words. However, researchers have found no direct relation between frequency levels and vocabulary acquisition. For example, Webb and Chang (2015) found no relation between word frequency levels and vocabulary learning. They asked 61 Taiwanese EFL learners to read and a series of 10 graded-reader books and listen to their companion audio CDs. They randomly selected 100 target words. After comparing the results of the pre-tests, post-tests and delayed-post-tests, they found a high gain in vocabulary by the participants even though no significant relation was found between frequency levels and vocabulary acquisition. However, they mentioned frequency as one of the effective factors. Later studies also supported the findings by Webb and Chang (2015). For example, Chang and Hu (2018) carefully replicated the study by Webb and Chang (2015). They divided 62 young adult learners into two groups of equal size. The participants in the first group (31) were high-level learners, and those in the second group (31), low level learners. The authors semi-randomly selected 100 target words, 31, 36, and 34 from the 1 000, 2 000, and 3 000+ highest frequency levels, respectively. After a series of reading and listening activities, they evaluated pre-test, post-test, and delayed post-test, whether the frequency levels had any effect on the acquisition of target words. They found that higher-level group learned 68%, 71%, and 69% while the lower-level group learned 20%, 19%, and 21% of the target words selected from the 1000, 2000, and 3000 frequency levels, respectively. Moreover, they found attrition rates of 7%, 11%, and 22% in the higher-level group and 31%, 34%, and 36% in the lower level groups in the respective frequency levels. They found no direct relation between word frequency level and vocabulary learning; moreover, they concluded that the vocabulary learning results in their study were possibly due to such factors as repetition, participants' reading and listening techniques, and after-reading activities.

3. Integrate new words with the old ones: the lexicon in humans is thought to be a web-like structure comprising interconnected links that create a network of associations (Aitchison, 1987, as cited in Schmitt & McCarthy, 1997, p. 241). For long-term vocabulary retention, teachers must help language learners to link newly acquired words to their previous knowledge and strengthen these links for better retrieval. Wessels (2011), in his proposed Vocabulary Quilt during reading, emphasizes the importance of the provision of opportunities for relating new words to learners' background knowledge in both pre-reading and during-reading phases as old knowledge, stored in background knowledge, assists learners to refine, expand and integrate word meanings into their long-term lexicon.

4. Providing several encounters with a word: "repetition is essential for vocabulary learning because there is so much to know about each word that one meeting with it is not sufficient to gain this information" (Nation, 2001, p. 74). Moreover, "repetition of foreign language forms promotes long-term retention"

(Ellis & Sinclair, 1996, p. 244). Repetition causes retrieval of either form or meaning. Each encounter with a word in a different context ended up with new links for that word in the lexical network. As discussed above, at least 10 encounters might be needed for THE successful incidental acquisition of a word (Jenkins, Stein, & Wysocki, 1984; Gathercole & Baddeley, 1990; Webb, 2007b; Bao, 2015). However, the repetition must also be systematic. It was found that spaced repetition was more effective than frequent repetition in one sitting (Baddeley, 1990; Bloom & Shuell, 1981; Dempster, 1987, as cited in Nation, 2001). The memory schedule Pimsleur (1967, as cited in Nation 2001) has been found to be an efficient repetition schedule for spaced repetition of word for long term acquisition. In Pimsleur's model, new words must be repeated at least eleven times within a fixed range of time starting at 5 seconds after the first encounter, followed by 25 seconds after the second, 2 minutes after the third, 10 minutes after the fourth, 1 hour after the fifth, 5 hours after the sixth, 1 day after the seventh, 5 days after the eighth, 25 days after the ninth, 4 months after the tenth, and 2 years after the eleventh encounter. Irrespective of the mode of presentation, repetition is an effective factor in vocabulary learning. For instance, Peters and Webb (2018) studied the role of repetition in audio-visual materials. They conducted two experiments with 63 intermediate level (B1-B2) Dutch-speaking EFL learners to find out if watching television could have any effect on form recognition, meaning recall, and meaning recognition. A full-length TV documentary program related to the course objectives of the participants, i.e., engineering subjects, with no subtitles was selected. They systematically selected 32 target words with frequency of occurrence ranging from 2 to 5. Data were collected via a pre-test-treatment-post-test design. Participants watched the full-length documentary after they had sat for the pre-test. After comparing the pre-test and immediate post-test results by ANOVA in order to find the most effective factors, they found that frequency of occurrence, or repetition, had a substantial impact on meaning recall, meaning recognition, and form recognition. Thus, the role of repetition, as an influential factor in vocabulary acquisition, should never be neglected or skipped by teachers or the researchers.

5. Promote deep level of processing: Craik and Lockhart (1972) showed that a deeper level of semantic processing would be more effective in acquiring a word. Hulstijn and Laufer (2001), in their study on levels of processing, found that retention was highest in the group who needed deeper semantic processing with the highest level of involvement owing to the need for deeper semantic processing. This technique is also key in this study; I discuss it in depth section 2.6.

6. Facilitate imaging and concreteness: Clark and Paivio (1991, as cited in Schmitt and McCarthy, 1997), in their theory of dual encoding, explain that the human lexicon is built from a network of the verbal and visual representations of different words. Accordingly, if learners try to learn a word through both channels, i.e., visual and verbal, retention will be better because a stronger link will be created in the lexicon for that word, although the effectiveness of this strategy could be challenged by the effect of context and multimedia materials. For example, in a multi-media context, the simultaneous presentation of text and

pictures, or videos may cause cognitive overload, which hinders learning (Mayer & Moreno, 2003). Moreover, Boers, Warren, Grimshaw, and Siyanova-Chanturia (2017) criticize most of the studies which support the effectiveness of dual coding theory for vocabulary uptake from texts for containing vocabulary items with concrete and imageable referents or for using previously created and established referents of the selected target words. Therefore, they emphasize the role of attention and posit that the "the attested benefits [of dual coding theory] lie with the amount of attention that multimodal clarification tend to attract" (Boers, Warren, Grimshaw, & Siyanova-Chanturia, 2017, p. 720). Finally, they conclude that dual coding theory is effective for vocabulary items do not have previously formed and established referents and images in the vocabulary learners' mind. A study by Shen (2010) supports the conclusions of Boers et.al., (2017). Shen (2010) tested the effect of dual encoding on the uptake of Chinese abstract and concrete words. She selected both abstract and concrete nouns for her study. Her participants were 45 males and females who had registered as beginners in a course on Chinese as a foreign language. The participants were divided randomly into two groups, namely, verbal-encoding only and verbal-encoding plus imagery. Shen administered two tests at the end of days 1 and 2 of her experiment immediately after instructions and tasks has been given and completed. In test 1, she asked the participants to select the correct words and in test 2 they were asked to write the meaning and the sound of the target words. Comparison of the test results showed that the verbal encoding plus the imagery method of instruction enhanced the uptake of the sound, shape, and meaning of abstract rather than concrete words. She concluded that the possible presence of previously stored images of the concrete nouns might have hindered acquisition of the target concrete nouns.

7. Dictionary work: When someone is asked to consider language learning, dictionaries are one of the first things mentioned. Dictionary work is considered an independent vocabulary acquisition strategy because it provides various ways of practicing words (Schmitt and McCarthy, 1997); however, choosing a good dictionary can probably help language learners achieve their goals faster and better. Hunt (2009) examined different types of monolingual and bilingual dictionaries and their effect on L1 and L2 vocabulary acquisition and retention. He found that although L1 is a good medium for conveying information and that bilingual dictionaries can better help understanding of the meaning, monolingual L2 dictionaries can be more fruitful for vocabulary acquisition and retention. Zou (2016) compared the effect of dictionary search on vocabulary learning with inferencing techniques. Participants were 104 intermediate English learners divided into two groups, namely reading comprehension and dictionary look-up, and reading comprehension and inferencing techniques. Immediate and delayed post-tests were administered after completion of the reading activities. The results indicated the superiority of dictionary search in learning target words. Zou (2016) found that the superiority of dictionary search can be explained by the degree of elaboration. She states that dictionaries provide an abundance of information on each word; therefore, a high degree of elaboration

provides the possibility for deeper word processing and adding more information about the looked-up words into the long-term memory. Moreover, she adds that as knowing a word includes knowing lots of relevant information, dictionaries can offer this complex and connected information for a word all at once, which in turn can create several links from the new words to existing background knowledge. Therefore, new knowledge is connected to old knowledge via various and informatively rich links.

8. Word unit analysis: All languages have a vast number of words. Thus, acquiring many words may be very exhausting. Teachers can help language learners to reduce this burden by teaching them the important affixes and roots of words in the target language (Sökmen, 1997). Bowers and Kirby (2010) investigated the effect of morphological instruction on vocabulary acquisition. They recruited 81 fourth and fifth graders for their study. Using a pre-test-treatment-posttest design, they taught and tested morphological knowledge of English words. During the treatment sessions, which comprised 20 50-minute instruction sessions, they taught spelling, morphemes, base words, affixes, and compounding words through problem-solving activities. To measure participants' meaning recognition and recall in both the pre- and post-tests, they selected 30 English words, which were to be taught during the treatment sessions. The analysis of the elicited data revealed that morphological instruction is a reliable tool for assisting language learners to acquire a large number of vocabulary items in the short term. It also assists them to acquire morphological awareness of the words in English. Although morphological instruction was effective for vocabulary acquisition, they reported that the effect of the instruction given did not extend to word families, which were not taught in their instruction sessions. In other words, although the participants acquired linguistic knowledge of English morphology, they were unable to decode the meaning of words other than those taught during the treatment sessions They concluded that morphological knowledge deepens the processes of acquiring a word. Thus, the quality of the mental representation of a word is enriched; this in turn lead to easier access to the already stored information about the word.

9. Mnemonic devices: these devices are used to assist memory during processing. They are verbal, visual, or a combination of both. The most famous mnemonic device, usually also cited as the most effective, is Atkinson's keyword method (Ellis & Beaton, 1993b; Sökmen, 1997). In this method, students first choose a word in their L1 that sounds like an L2 target word and then try to form a visual picture which is a mixture of both words (Sökmen, 1997; Ellis & Beaton, 1993b; Nation, 1990). Brown and Perry (1991) compared three mnemonic devices, i.e., keyword, semantic, and key-word-semantic. Their participants were 60 Egyptian Arabic-speakers distributed in 6 ESL intact classes. In a nonequivalent control group design, they assigned the participants to either the experimental or control group. They selected 31 nouns and verbs as their target words in the study. They asked an Arabic native speakers to select the Egyptian colloquial Arabic keywords, which could be nouns or verbs, and a sound-alike English target word of at least in one syllable, for each keyword. During the treatment

sessions, the keyword group received the target words, their definitions, and the keywords. The participants in the keyword-semantic group received the target words, their definitions, the keywords, example sentences, and questions. After the treatment, they administered a 40-item, four-choice multiple item test containing the 31 target words. Analysis of the test data showed that the most effective mnemonic device was the keyword-semantic method. They suggest that the effectiveness of this method can be explained by the Depth of Processing hypothesis.

> That is, first, information processed at the semantic level produces better memory traces than that processed at acoustical and visual levels; and second, when elaboration occurs at a number of levels, memory traces are even stronger. If assumptions that recognition tasks measure information stored in memory and cued-recall tasks reflect facility in retrieving information from memory are warranted, then the conclusion can be made that the combination of these methods produced both stronger memory traces and better retrieval paths than if used alone. (Brown & Perry, 1991, pp. 665-666).

However, Ellis and Beaton (1993b) found the keyword method to be superior for gaining receptive vocabulary knowledge and repetition to be the most effective factor in facilitating the acquisition of vocabulary productive knowledge.

10. Semantic elaboration: in these types of activities, the teacher seeks to promote the formation of associations and thereby build up students' semantic networks (Machalias, 1991, as cited in Schmitt & McCarthy, 1997, p. 249). Among the best known types of these activities are semantic feature analysis, in which the meaning components of a word are analyzed, pictorial schemata, which refers to creating grids or diagrams to present lexical ordering, and semantic mapping, which refers to digging deeply into the associations of a word and presenting these in a diagram in order of relevance. Semantic elaboration techniques help learners "distinguish differences in meaning and organizing words" (Schmitt & McCarthy, 1997, pp. 251-2). Taevs, Dahmani, Zatorre, and Bohbot (2010) state that semantic elaboration affects auditory and visual memory. Therefore, in using semantic elaboration techniques to teach vocabulary acquisition by means that deeper semantic processing is possible because it involves the storing of audio-visual as well as lexical data about words. Therefore, the new knowledge is related to the pre-existing knowledge through strong and informatively rich links, which in turn makes it easier to access the stored words (Thuy, 2013).

11. Collocations and lexical phrases: Schmitt (2010a) terms these formulaic sequences. Sökmen (1997) claimed that because collocational relationships create long-run links and their traces are long-lasting, providing students with an opportunity for practicing them would seem a promising method of learning. Boers, Eyckmans, Kappel, Demecheleer, and Stengers (2006) found that more knowledge of formulaic sequences improved language learners oral proficiency. They investigated how knowledge and the use of formulaic sequences could assist language learners to increase their proficiency in L2 use. They also tested the effect of awareness on the acquisition of formulaic sequences. 32 participants, majoring in English, were randomly divided into either the experimental or

control group. They received 22 hours of teaching, including the use of several authentic reading and listening materials. The experimental group participants were informed about standard word combinations while they were interacting with the reading and listening materials. The control group participants were taught using the traditional grammar-vocabulary dichotomy. After their respective treatments, the oral proficiency of the participants was evaluated by interviews and the rate of their uses of formulaic sequences were measured by judges blinded to group allocation. Analysis of the data revealed a positive correlation between oral proficiency and knowledge of formulaic sequences. Moreover, the researchers reported that augmenting the language learners' awareness of formulaic sequences helped them enhance their oral proficiency and language to the extent that they came across as proficient speakers. Boers, Eyckmans, Kappel, Stengers, and Demecheleer (2006, p. 256) state that according to cognitive linguists "the category of formulaic sequences contains segments that are much less arbitrary than used to be assumed, and [their studies] therefore lend themselves well to insightful learning and mnemonic strategies". Overall, cognitive linguists posit that by encouraging mnemonic strategies, such as imagination, alliteration and assonance, formulaic sequences boost acquisition of these specific word combinations.

12. Vocabulary learning strategies: Schmitt (1997) provides a taxonomy of vocabulary learning strategies, based on "Oxford's (1999) system and the discovery/consolidation distinction" (p. 204). His taxonomy is categorized into four groups of distinctively different strategies: social strategies, referring to the use of interaction for acquiring a word, memory strategies, referring to any activity that relates new material to existing knowledge, cognitive strategies, or activities in which learners manipulate the target language, and metacognitive strategies, referring to processes that monitor vocabulary acquisition processes. According to Gu and Johnson (1996) and Gu (2003), the type of strategy for learning vocabulary items that language learners choose highly depends on factors such as the person, context and task at hand. Gu (2003) considers a vocabulary learning task as a problem-solving activity which is interacted by a learner or, say, a person who has myriad attributions and qualities. A person chooses a strategy by considering internal factors, such as motivation, anxiety, proficiency, background knowledge, etc., contextual factors, such as learning culture, quality of input and output, etc., and task-related factors, such as difficulty, time, complexity, etc. Therefore, he recommends, first, that language teachers should be aware of the importance of vocabulary learning strategies in vocabulary learning, and second, that before implementing any vocabulary learning strategies, language teachers must search for the most effective vocabulary learning strategies and tasks for specific groups of language learners and their cultures and context of learning.

2.4.4 How many words does one need to learn?

Hulstijn (2001) writes that determining how many words one needs to learn is unlikely; however, many empirical findings have suggested answers to this

question. Based on the literature on L2 learning, Hulstijn (2001) nevertheless proposes at least 5 000 base words as the minimal learning target for receptive knowledge. However, he argues that this estimate is not enough since to reaching the 95 percent level of word knowledge required to read different non-specialist texts requires at least 10 000 base words.

Milton (2009) observes that reading for study purposes and reading for pleasure are different tasks and that Nation's (2006) estimate pertains to reading for pleasure. Milton's finding that for unassisted comprehension of a general text, either written or spoken, 98% of running words in the text must be known means that learner must acquire the previously mentioned amount of word-family knowledge. In turn, Laufer and Sim (1985, as cited in Milton, 2009) estimated that 95% knowledge of running words needed for unassisted comprehension of academic texts. Milton (2009) concludes that the estimate for unassisted comprehension of a text is "probably 95% or more" (p. 51). He adds that with knowledge of at least 2 000 word-families learners can understand the general idea of the text. "Very broadly, this suggests, not surprisingly, that the more vocabulary the learners know, the more they think they understand both in reading and in speech" (Milton, 2009, p. 53). However, Milton's conclusion is highly referable to English.

2.4.5 How is knowing a word studied?

Perusal of the previous studies and literature reviews suggests that vocabulary acquisition has largely been studied with respect to specific factors and by tapping into different aspects of word knowledge. Thus, knowing a word and vocabulary acquisition has been studied by considering factors such as size of vocabulary, depth of vocabulary knowledge, the incremental nature of vocabulary learning, the importance of word forms, vocabulary acquisition in L2 compared to L1, engagement with words, phrasal vocabulary, exploring methods for increasing the number of exposures to vocabulary items, effective methods of repetition, aspects of lexical knowledge, incidental vocabulary knowledge, number of exposures required for incidental vocabulary knowledge, incidental learning from listening, extensive reading and context, glossing, and how to mix incidental and intentional vocabulary learning and teaching techniques for efficient vocabulary learning (Reed, 2004; Schmitt, 2008; Laufer, 2009).

Barcroft (2004) summarizes previous research foci into 10 categories. He reports that almost all the major studies that have been conducted on knowing and learning a word in second language acquisition can be assigned to one of those 10 categories. In the following table, Barcroft (2004) presents a summary of the areas of research and some major findings from each perspective of vocabulary research.

TABLE 4 Research Areas Related to L2 Vocabulary Acquisition (Adapted from Barcroft, 2004, p. 202)

1. Incidental vocabulary learning	• Learners pickup new words while reading text without being instructed to do so (Nagy, Herman, & Anderson, 1985). • Learners frequently rely on inferencing strategies when dealing with unknown words in texts (Paribakht & Wesche, 1990). • Reading proficiency and topic familiarity affect word gain (Pulido, 2003). `
2. Lexical requirements for comprehension [*or word coverage knowledge for unassisted text comprehension*]	• Bribois (1995) demonstrated the critical role of L2 vocabulary knowledge for successful L2 reading comprehension. • Nation and Waring (1997) calculated that with 2000 words learners can reach 80% text comprehension. They also posited that approximately 3000 words of high frequency in a language are an "immediate high priority".
3. Input enhancement and text-based factors [*or how to manipulate target words as inputs during learning*]	• Davis (1989) found positive effects for marginal glossing on L2 reading. • Hulstijn, Hollander, and Greidanus (1996) found incidental vocabulary learning was facilitated by increasing word frequency in a text, bilingual dictionary use (as compared to a control) and definition in marginal glosses (as compared to dictionary use • Typological enhancement productive learning of enhanced L2 word while decreasing productive learning of unenhanced L2 words (Barcroft, 2003).
4. Vocabulary learning strategies	• Atkinson and Raugh (1975) demonstrated the effectiveness of Keyword Method. • Ahmed (1989) found that more successful vocabulary learners tend to utilize a larger and more varied repertoire of vocabulary learning strategies, are aware of their learning, and are more aware of the semantic relationships between new and previously learned words. • Schmitt (1997) developed a taxonomy of 50 vocabulary learning strategies based on those used to infer meaning and those used to consolidate words.
5. Combined incidental and direct vocabulary instruction	• Combining incidentally oriented and direct methods of L2 vocabulary instruction can be more effective than using indirect, incidentally oriented instruction alone (Paribakht & Wesche, 1997).
6. Methods of direct instruction	• Prince (1996) found translation-based L2 vocabulary learning to be more effective than learning vocabulary in the context of sentences. • Barcroft (1998, 2000, 2004) found negative effects for writing new words in sentence on productive L2 vocabulary learning.
7. Word-based determination of learnability [*or what factors are in words that make them difficult to learn*]	• Ellis and Beaton (1995) found that longer words and L2 words less phonologically similar to L1 words were more difficult to learn. • Laufer (1997) found that "deceptive transparency" (incorrectly inferring the meaning of an expression based on the knowledge of the words within the expression) can make it more difficult to learn L2 words.
8. Bilingual mental lexicon [*study of bilingual people word knowledge*]	• Potter, So, Von Eckardt and Feldman (1984) compared two models: *concept mediation*, positing direct connections between L2 words and concepts; and *word association*, positing that L2 words are connected to concepts through L1 words. Their findings were interpreted as support for concept mediation.

	• Subsequent research on proficient bilinguals suggests that lexical and semantic information in L1 is activated during comprehension and production in L2 (review in Kroll & Sunderman, 2003). • Kroll, Michael, Tokowicz, and Dufour (2002) found that performance on tasks in L1 can be affected by L2 proficiency level: Advanced bilinguals named L1 words faster than did beginning L2 learners of the same L1.
9. Receptive versus productive vocabulary knowledge	• Estimates have been made that receptive vocabulary knowledge is twice as large as productive vocabulary knowledge (see Marton, 1977). • The productive-receptive distinction may exist at the level of testing and not in the learner's mind. Melka (1997, p. 101-102) noted: "it is certainly not clear whether [reception] and [production] ought to be considered as two separate systems depended on each other, or rather as one unique system (one lexical store) used in two different ways, receptively or productively."
10. Lexical input processing [*or how a target word for learning is processed, analyzed, and added to the mental lexicon of a learner*]	• Focusing extensively on the meaning of new L2 words sometimes can inhibit learning the formal properties of those words (Barcroft, 2002). • Forced output (e.g., requiring learners to write new L2 words in sentences) can decrease learning the formal properties of L2 words (Barcroft, 2000).

Drawing on table 4, I used glosses and pre-teaching techniques in this study. Previous studies have mostly reported a positive effect of glosses, especially L1 glosses (Rott, Williams, & Cameron, 2002), and pre-teaching new words on the acquisition of vocabulary items (Nation, 2002; Rott & Williams, 2003; Laufer, 2006; Yoshii, 2006; Lin & Huang, 2008; Schmitt, 2010b; Xu, 2010). Nation (2002) mentions that glosses make texts more comprehensible and that they prevent wrong guesses. Rott and Williams (2003) also found that reading plus glosses can be effective for vocabulary acquisition because "a) glosses trigger[…] a search for concrete meaning and firm form-meaning mapping; b) a lack of glosses correspond[…] with global text processing, skipping of words and shallow meaning mapping; c) multiple encounters help[…] +gloss readers to gain semantic information" (p. 45).

Regarding the pre-teaching technique, Schmitt (2010) adds that using explicit vocabulary learning techniques increases the chance of uptake by focusing learners' attention on the vocabulary items of interest. Laufer (2006) studied the effect of a pre-teaching technique on vocabulary acquisition. She recruited 158 high school English learners and categorized them into two groups. One group learned the target words incidentally by reading a text for comprehension purposes. The other group were pre-taught the same vocabulary items before reading the text. Comparison of their post-test scores revealed a significant gain in vocabulary acquisition by the group who were pre-taught the target words.

In the present study, L1 definitions were also pre-taught and glossed for three obvious reasons. First, more comprehensible input leads to more effective L2 vocabulary acquisition (Barcroft, 2004). Second, Hong (2010), in his review of

the role of glosses in the language learning literature, concludes there is no significance difference between the effect of L1 and L2 glosses in vocabulary acquisition. Third, Rott, Williams, and Cameron (2002) found that the mental effort induced by L1 glosses is effective in the acquisition and long-term retention of target words both productively and receptively.

2.5 What does make it hard to acquire a word?

Many factors can make acquiring L2 words difficult. Numerous formal, phonological, and psychological factors determine the level of difficulty of a word (Nation, 1990; Ellis & Beaton, 1993a; Laufer, 1997). Ellis and Beaton (1993a), and Laufer (1997) propose the following factors as the most influential in the determination of word difficulty:

1. *Familiarity of phonological features and pronounceability*: there may be problems in acquiring a word when language learners encounter a new phonological or articulatory feature in the target language. For example, the distinction between the sounds /u/ and /y/ in Finnish may affect the acquisition by Persian speakers of Finnish words (e.g., tuuli – tyyli) distinguished by this feature. Clearly, the pronunciation and articulation of a word containing sounds that are not present in the new language learner's L1 inventory can be difficult. The learner may refuse to use such words, rather like Celce-Murcia's bilingual daughter, who refused to use French words that she found difficult to pronounce (Schmitt & McCarthy, 1997). Ellis and Beaton (1993a) conclude that "1. The less the overlap between the feature set of the native and the foreign language, the harder it will be for the FL learner to learn to speak that language; 2. The less the overlap between the feature set of the native and the foreign word, the harder it will be for the FL learner to learn that word" (p. 561). Rodgers (1969, as cited in Ellis and Beaton, 1993a) found that English learners of Russian learned words with difficult pronunciations very late.

2. *Orthography*: the degree of sound-script correspondence may also influence the level of difficulty in learning a word. Moreover, in a language with a writing system different from learner's L1, the degree of difficulty in acquiring a word is even higher. Ellis and Beaton (1993a) add that sequential letter probability can also affect the degree of difficulty in learning a word; "Thus, the learning of the orthography of FL words may be determined by the degree to which the sequential letter probabilities match those of the native language […] the degree to which a particular FL word accords with the orthographic patterns of the native language may affect its ease of learning." (Ellis and Beaton, 1993a, p. 567-568).

3. *Length*: Laufer (1997) posits that longer words are harder to learn. Her view accords with Zipf's law. According to Zipf, the more frequent words in a language tend to be shorter in both syllables and letters (Milton, 2009). Ellis and Beaton (1993a) argue that word length increases difficulty because there is more

to remember and there is also broader scope for orthographic and phonotactic variation that may induce further errors.

4. *Inflectional complexity*: Laufer (1997) states that "irregularity of plural, gender of inanimate nouns, and noun cases make an item more difficult to learn […] since the learning load caused by the multiplicity of forms is greater" (p. 145).

5. *Derivational complexity*: in a language, where there is, for instance, morpheme combinations generate a multiplicity of meanings, learning a word can be is difficult (Laufer, 1997). For example, in English, learners must know that 'outline' is correct but 'preline' is not. Or they must be aware that 'outline' does not mean 'out of line' in the way that the 're' in 'review' means 're: again + view'.

6. *Synformy*: Laufer, introducing the term synformy, defines it as "similarity of lexical forms" (1997, p. 146). According to Laufer (1988, as cited in Kocić, 2008, p. 52), there are 10 different classes of synformy. Category 1 comprises words with the same roots that are productive with different suffixes in current English, such as successive/ successful. Category 2 contains words with the same roots that have different suffixes but which are not in productive use in present-day English, for example, credulous/credible. Category 3 includes synforms that differ in a suffix which is present in only one of them, for instance, historical/historic. Category 4 contains synforms that have the same roots but different prefixes and are not in productive use in modern English, such as resumption/consumption/assumption. Category 5 contains synforms that are different in a prefix, which exists in only one of them, like commission/mission. Category 6 comprises synforms hose phonemes are identical except for one vowel or one diphthong in the same syllable or position, for instance, affect/effect. Category 7 consists of synforms that differ in one vowel which is present in only one of them, for example, quite/quiet. The category 8 synforms have identical phonemes but differ in only one consonant, like extend/extent. The synforms in category 9 are identical in all phonemes except for one consonant that is present in only one of them, such as, stimulate/simulate. Category 10 contains synforms that have identical consonants but different vowels, for instance, menial/manual. Laufer (1997) believes that synformy causes difficulty in learning a word because learners confuse words with an almost similar sound or look. Schmitt (2010a) also reports synformy as one of the most important factors affecting the learning of a word form.

7. *Word class*: certain grammatical categories are easier to acquire than other types. For example, nouns are the easiest followed by adjectives and verbs. Adverbs are believed to be the hardest class of words to learn (Ellis & Beaton, 1993a). For example, Philips (1981, as cited in Schmitt & McCarthy, 1997, p. 148) found that nouns are learned more easily than other classes of words.

8. *Abstractness*: Ellis and Beaton (1993a) call this feature the imageability of a word, and as the name suggests it means the extent to which a word can generate a mental image. Laufer (1997) states that abstract words, like the word 'lie', are much harder to learn than concrete words, like the word 'bread'.

9. *Specificity and register limitation*: Blum and Levenston (1978, as cited in Laufer, 1997) report that non-native learners use more general terms in their

writing compared to native writers who try to use more specific terms. Laufer (1997) suggests that the main reason is the unfamiliarity of learners with the register limitations of the L2 words. She concludes that words with very limited registers, for example *cordillera*, are difficult to learn.

10. *Idiomaticity*:

> We think of a locution or manner of speaking as idiomatic if it is assigned an interpretation by the speech community but if somebody who merely knew the grammar and the vocabulary of the language could not, by virtue of that knowledge alone, know (i) how to say it, or (ii) what it means, or (iii) whether it is a conventional thing to say. Put differently, an idiomatic expression or construction is something a language user could fail to know while knowing everything else in the language. (Fillmore, Kay, & O'Connor, 1988, p. 504)

On idiomatic expressions, Dagut and Laufer (as cited in Laufer, 1997) found that L2 learners tend to use one-word verbs rather than phrasal verbs, or in this case, idiomatic expressions. Thus, in the process of language learning, for a learner using "*decide* would be easier that *make up one's mind* "(Laufer, 1997, p. 151). Therefore, because of the obstacles and pitfalls that idioms present for language learners, such as learning load, idiomaticity increases the difficulty of learning a word (Laufer, 1997).

11. *Multiple meaning*: in linguistics, this feature is called polysemy or homonymy (Yule, 2014). "Empirical evidence is available to illustrate the difficulty learners have with polysemy and homonymy" (Laufer, 1997, p.152). For example, Bensoussan and Laufer (as cited in Laufer, 1997) found that comprehending words with multiple meanings was more difficult and caused more errors in understanding the text.

12. *Word frequency*: it is believed that more frequent words are easier to learn because frequent exposures increase the chance of learning (Ellis & Beaton, 1993a). Reversing Zipf's law it can be assumed words with low frequency are difficult to learn (Milton, 2009). Word frequency can be related to word register. Technical words, which have very limited registers, are not used by everyone and thus have low frequency. This leads to the conclusion that because they have low frequency and limited registers such words hard to acquire.

13. *Degree of meaningfulness*: the last, but not least, factor in determining the difficulty of a word is the extent to which a word is associated with its meaning (Ellis & Beaton, 1993a). They report that "When both stimulus [word] and response [meaning] are more meaningful, there is a greater chance of forging associations between them" (p. 567). Therefore, when a word has elements of form, shape or sound, etc. that invite more associations with its meaning, that word is easier to acquire than words with a small number of such elements. For example, onomatopoeic words, like *bang,* or *ring*, are easier to learn than an abstract word like *lie*.

2.6 Levels of involvement load hypothesis (ILH):

In the vocabulary learning literature, it has been suggested that the level of involvement load is a reliable index that can predict the efficacy rate of vocabulary learning tasks (Hulstijn & Laufer, 2001). The ILH has been investigated by many researchers, especially in studies on incidental vocabulary acquisition; and despite mixed results on its success in predicting effectiveness (Keating, 2008; Yaqubi, Rayati, & Gorgi, 2010; Kim, 2011; Tahmasbi & Farvardin, 2017; Zou, 2017), it is commonly considered an important factor in vocabulary learning tasks (Huang, Willson, & Eslami, 2010). It can, therefore, be recommended that vocabulary learning tasks should be evaluated before their application in the classroom, especially if they are designed for incidental vocabulary acquisition. In the present study, the effect of task-induced levels of involvement load was evaluated to investigate the efficacy of digital game tasks in vocabulary acquisition. However, before moving on to the materials drawing on the involvement load hypothesis that are used in this study, it is necessary to revisit and discuss its history.

The ILH has its origin in the memory studies by Craik and Lockhart (1972) and Craik and Tulving (1975). In a seminal paper, Craik and Lockhart (1972) introduced the concept 'depth of processing'. They argued that "memory trace is a by-product of perceptual analysis. The persistence of that trace depends on how deep the stimulus has been analyzed" (Craik & Lockhart, 1972, p. 671). They continued by stating that more elaboration was associated with longer persistence of the trace. In other words, "rich (qualitative) and numerous (quantitative) associations with existing knowledge […] increases the chance that the new information will be retained" (Hulstijn & Laufer, 2001, p. 541). However, Craik and Tulving (1975) posit that the notion of depth differs from that of elaboration because they believed that if the depth and elaboration are the same, they had provided nothing new.

A new stimulus is analyzed in different stages. For example, in the early stages, its shape, lines, sound etc., which are easy to recognize, are noticed and analyzed while later stages are concerned with connecting the new input with previous knowledge or with recognizing new patterns and extracting meaning. "This conception of series or hierarchy of processing stages is often referred to as depths of processing" (Craik & Lockhart, 1972, p. 675). Craik and Tulving (1975, p. 291) hold that elaboration can happen at every stage of the analysis because "the basic core of the events can be elaborated in different ways". In sum, "memory trace persistence is a function of the depth of analysis, with deeper levels of analysis associated with more elaborate, longer lasting, and stronger traces" (Craik & Lockhart, 1972, p. 675).

Although Craik and Tulving's (1975) empirical study argued for the importance of the existence of levels of processing and of the positive effects of degrees of elaboration during encoding processes on retention, it was challenged by other scholars. The major criticisms were contained the following two

questions: 1) What, exactly, are the levels of processing? and 2) How can we know which level is deeper than another? (Baddeley, 1978; Eysenck, 1978; Nelson, 1977, as cited in Hulstijn & Laufer, 2001; Hulstijn, 2001).

Craik and Lockhart were unable to present any operational definition for the levels of processing; moreover, they could not provide any method to measure the depth of processing levels. However, cognitive psychologists concede that processing a lexical entry more deeply and more elaborately would lead to better retention (Hulstijn & Laufer, 2001). In practice, and for word learning, it means that if learners pay more attention to different aspects of a word, such as its pronunciation, orthography, meaning, grammatical category, and semantic relations with other words, they will learn it better than when they attend only, for example, to pronunciation and orthography (Hulstijn & Laufer, 2001; Hulstijn, 2001).

Laufer and Hulstijn (2001), in developing their hypothesis, took cognizance of the importance of depth of processing, levels of processing, and richness of elaboration in retention. They also considered the role of motivation, since they believed that humans are not mere information-processing entities; they also have motives and emotions that may affect their information-processing procedure (p. 6). Putting all these considerations together, they developed a motivational-cognitive construct of involvement named the 'levels of involvement load hypothesis for L2 vocabulary acquisition (Hulstijn & Laufer, 2001).

Their construct of involvement consists of three components: need, search, and evaluation. Accordingly, they assumed that the "retention of words when processed incidentally is conditional upon the following factors in a task: need, search, and evaluation" (Laufer & Hulstijn, 2001, p. 14). That is, different combinations of need, search, and evaluation define the processing load of different word learning tasks. They call this combination 'involvement'.

They further define the components of involvement. Need is the motivational non-cognitive part of the hypothesis that refers to the extent to which a learner feels that an unknown word must be learned (Laufer & Hulstijn, 2001; Schmitt, 2008). The need is moderate if it is provoked by an external agent and strong if it is provoked by the learners and their internal factors. For example, when a learner meets a word which is important for the comprehension of a text, he feels that there is a need to learn it. If the learner is asked by the teacher to learn that word, the need is moderate. If the learner wants to learn the word to expand his lexicon, the need is strong (Laufer & Hulstijn, 2001; Hulstijn & Laufer, 2001; Schmitt, 2008).

Search and evaluation are the two cognitive components of the hypothesis that concern form-meaning relationships (Hulstijn & Laufer, 2001). Search is a learner's attempt at finding either the meaning of an unknown L2 word or the form of a concept in L2 by consulting a dictionary or another authority, e.g., a teacher (Hulstijn & Laufer, 2001). The model does not state whether search has levels; in their examples, search is dichotomous: it either exists or does not exist (Laufer & Hulstijn, 2001; Hulstijn & Laufer, 2001).

Finally, they define evaluation as any internal or external comparison performed for a word. For example, if a word is compared to its surrounding words or if one meaning of a word is compared to other meanings of that word to see which one fits the context best, the learner has performed evaluation. The evaluation is moderate if the learner tries to recognize a word form from among other different word forms or to compare senses of a word. The evaluation is strong if the learner needs to decide about additional words which will combine with the new word in the original context (Laufer & Hulstijn, 2001, p. 15). For example, when in writing a composition the learner must find a form for an L2 concept in a dictionary and use it in a sentence, evaluation is high. In terms of indexing, according to this model, the maximum index that a vocabulary learning task can induce is 5.

Overall, the authors hypothesize that a task that demands greater involvement provides better possibilities for L2 word acquisition than a task demanding a lower level of involvement (Laufer & Hulstijn, 2001). To test their hypothesis, Hulstijn and Laufer (2001) compared 225 learners, assigned to three groups, in their short- and long-term retention of ten unfamiliar words. Controlling for time, they compared participants' rates of incidental acquisition of the target vocabulary items in reading comprehension groups, comprehension plus filling in the target words, and composition-writing with the target words. As they expected, the composition-writing groups whose task required the highest level of involvement, retained the target words better than other groups.

The existence and effectiveness of levels of involvement in vocabulary acquisition have also been observed by other researchers. Jing and Jianbin (2009) applied the hypothesis to vocabulary acquisition during three listening tasks differing in the level of involvement. They gave the tasks to three parallel classes each containing 29 non-English major students. Using the Kruskal-Willis test, they tested the correlation between task involvement load and retention. Their results demonstrated the validity of the involvement load hypothesis in retention, and they concluded that higher levels of involvement boost the retention of vocabulary items. Yaqubi, Rayati, and Gorgi (2010) tested the effect of level of involvement load on L2 vocabulary retention in an academic EFL context. Participants were 60 EFL learners divided into three groups, two of were tasked with acquiring vocabulary items via two input-oriented tasks with involvement load indices of three and two, respectively. The third group completed an output-oriented task with an involvement load index of three. The third group showed the highest level of involvement. They also concluded that task type can influence the effect of involvement load and hence that the construct of evaluation needs reconsideration.

Although involvement load has been shown to be influential in word learning task outcomes, Nation and Webb (2010) note that "the involvement load hypothesis does not include many features [e.g., repetition and time on task] that other researches have been shown to be important when designing vocabulary teaching techniques" (p. 7). Moreover, other researchers (Folse, 2006; Keating, 2008; ; Martínez-Fernández, 2008; Kim, 2010; Jahangiri & Alipour, 2014; Zou,

2017) have criticized the original method of quantifying task-induced involvement load proposed by Hulstijn and Laufer (2001), along with its components such as need, search and evaluation, arguing that due to lack of precision in measuring the components, especially evaluation, the predicative power of the current involvement load indexing method is not reliable. Therefore, Nation and Webb (2010) conclude that the index is not a good instrument for designing word learning tasks and vocabulary teaching techniques, and that a more elaborate set of criteria is needed. They introduce their own checklist which they call 'Technique Feature Analysis'. In it, they have expanded the three dimensions of involvement and combine these with other factors that are influential in indexing the involvement load of a word learning task. Their checklist comprises five main categories: Motivation (which also contains the notion of need), Noticing, Retrieval (which also covers the concept of search), Generation and Retention. The maximum checklist score is 18, one point for each of the 18 ILH index criteria. Each criterion is scored 0 or 1. To test their checklist, they have analyzed many word learning activities and identified many differences in other indices when compared to the ILH index. For example, multiple-choice text and word cards have an index value of 3 in the IHL and of 3 and 11, respectively, in technique feature analysis (Nation & Webb, 2010, p. 14). The superior accuracy of technique feature analysis, as an index of the involvement loads of L2 vocabulary learning tasks, has been investigated and argued for by, e.g., Hu and Nassaji (2016), Chaharlang and Farvardin, 2018, Gohar, Rahmanian, and Soleimani (2018), and Zou and Xie (2018). In this study, I also used the checklist to construct an index of task load in order to more precisely evaluate the effect of different levels of involvement load. The checklist is presented in the following table.

TABLE 5 A checklist for technique feature analysis (Adapted from Nation & Webb, 2010, p. 7)

Criteria	Scores
Motivation	0 1
Is there a clear vocabulary learning goal?	0 1
Does the activity motivate learning?	0 1
Do the learners select the words?	0 1
Noticing	
Does the activity focus attention on the target words?	0 1
Does the activity rise awareness of new vocabulary learning?	0 1
Does the activity involve negotiation?	0 1
Retrieval	
Does the activity involve retrieval of the word?	0 1
Is it productive retrieval?	0 1
Is it recall?	0 1
Are there multiple retrievals of each word?	0 1
Is there spacing between retrievals?	0 1
Generation	
Does the activity involve generative use?	0 1
Is it productive?	0 1
Is there a marked change that involves the use of other words?	0 1

Retention	
Does the activity ensure successful linking of form and meaning?	0 1
Does the activity involve instantiation?	0 1
Does the activity involve imaging?	0 1
Does the activity avoid interference?	0 1
Maximum score	**18**

2.7 Instruments Background

To arrive at a more precise evaluation and to study the effect of task-induced levels of involvement load on the acquisition of target words in greater depth, I used a specific instrument and design to collect the qualitative data. By using introspective methods (Ericsson & Simon, 1987), such as concurrent think aloud, in which participants report their mental events while they are engaging with a task, especially problem solving tasks, and an exit interview after task completion, I was able to investigate participants' minds and gain a clearer picture of the effect. Furthermore, after considering the different methods of collecting introspection data, such as think aloud, talk aloud, concurrent think aloud, retrospective think aloud (Ericsson & Simon, 1987), self-report, self-revelation, and self-observation (Cohen, 1987), I selected concurrent think aloud. In the following sections, to explain my choice, I briefly review the verbal report as a method of collecting data on vocabulary acquisition.

2.7.1 Think Aloud

The "think aloud" or "protocol analysis" was originally developed by Newell and Simon in 1972 from introspection, which had been developed to explore problem-solving strategies (Someren, Barnard, & Sandberg, 1994; Katalin, 2000). However, Charters (2003, p. 69) argues that think aloud can also be traced to Vygotsky's concept of "inner speech", in which Vygotsky had theorized that inner speech in adults' verbalization of thought evolved from two sources: from "egocentric speech of toddler monologues" and from "a form of thinking aloud with the goal of solving problems". Think aloud has been widely used in psychology, especially in the field of cognitive science, to investigate mental processes that are not easy for researchers to monitor, analyze and explore (Yoshida, 2008). As Ericsson and Simon (1987, p. 24) state "after a long period of studying human performance and abilities, research in psychology is now seeking to understand the underlying cognitive processes". Like introspection, in which events in progress was observed as stream of consciousness, in think aloud, the research subjects are asked to verbalize what they are thinking and whatever is passing through their minds while they are busy with a task. They must try to think aloud as if they are alone in their private rooms and are busy with a task (Yoshida, 2008).

2.7.1.1 Theoretical Background and Validity

Although think-aloud and the use of raw data collected through verbalization fell into disrepute during the era of behaviorism, the emergence of cognitivism in the 1960s, and especially the information-processing framework, thanks to the work of Andrew Ericsson and Herbert Simon, elevated the status of think aloud to a respectable scientific method of data collection, (Saravia, 1995; Charters, 2003). Simon and Ericsson (1987), in their seminal paper "*verbal reports on thinking*", defended the use of verbal reports and argued that the method could be considered scientific. They argued that information processing is the core of human cognition. That is, cognitive processes can be viewed as "a sequence of internal states" that are arranged in a specific order and can be filled in with information (Ericsson & Simon, 1987, p. 25; Saravia, 1995). They posit that the information is stored in several types of memory[2] that have different capacities and can be separately accessed. Thus, when a research subject starts to think aloud, he verbalizes thoughts or pieces of information that are in one of these stores, but which one? Ericsson and Simon claim that information in short-term memory (STM) can be verbalized because the information is available for further processing and is attended to. They call this type of information, or thought, "heeded information" (Ericsson & Simon, 1987, p. 32). They explain that heeded information is also recent information because, as in the information-processing framework, the information must be transferred from the long-term memory to short-term memory to be ready for verbalization. Moreover, they consider that only the information in STM is attended to and processable.

Ericsson and Simon Point out that "the sequence of states, i.e., the information contained in attention and STM, remains the same with the verbal reports as it would be when the cognitive processes proceed silently" (Ericsson & Simon, 1987, p. 27). In other words, they discovered that although thinking aloud affects and increases the duration of problem-solving tasks, it does not affect, or alter the online processes that are running during the subject's engagement with the problem-solving task. They put their theory into practice by gathering the same data via both think aloud and talk aloud, i.e., when the subject simply verbalizes silent or inner speech. They found that the processes were almost the same in both research settings and, as they expected, the only distinctive factor was duration, which was longer for the think-aloud group (Ericsson & Simon, 1980, 1987). They conclude that the data collected via think aloud can be considered scientifically valid.

---

[2] These memories are: 1) "sensory stores of very short duration", 2) a short-term memory (STM), which has a small capacity, and 3) a long-term memory (LTM), which has the largest capacity and can store a lot of information permanently (Ericsson & Simon, 1987, p. 26).

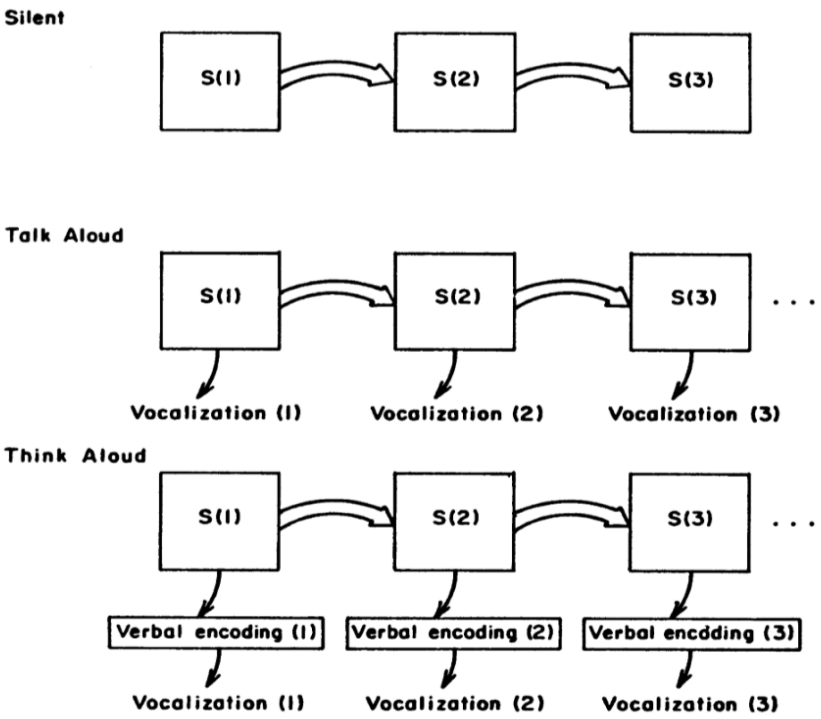


FIGURE 3 "the state of heeded information in a cognitive process and their relation to verbalization under three different conditions" (Ericsson & Simon, 1987, p. 33).

2.7.1.2 The Think-Aloud Protocols and Types

Ericsson and Simon (1980) emphasize that the "interval between the moment of acquisition and the moment of recall […] is an important consideration in classifying verbalization procedures" (p. 218). They broadly classify the think-aloud protocols, or methods of verbalization, into two types: *concurrent verbalization* and *retrospective verbalization*. In concurrent verbalization, the research subject is asked to verbalize the online processes and information that he or she is attending to while completing a task. On the other hand, if the subject is asked to verbalize the cognitive processes and the heeded information after he or she completed the task, the procedure will be labeled as retrospective verbalization (Ericsson & Simon, 1980; Yoshida, 2008).

Leow and Morgan-Short (2004), following Ericsson and Simon's categorization, introduced two other types of verbal reports, namely *metalinguistic* and *non-metalinguistic* verbalization.

> In the metalinguistic verbalization, the researcher may ask for specific information (e.g., reasoning or explanation), and learners provide a metacognitive report on what they think their processes are. In non-metalinguistic verbalization, learners are focused on the task with the think aloud secondary and only voice their thoughts without explaining them. (Leow & Morgan-Short, 2004, p. 36).

Although these are possible, valid, and experienced methods for collecting verbalization data, retrospective think aloud has been criticized for the probable effects of memory constraints. Accordingly, concurrent non-metalinguistic think aloud is recommended for collecting data on cognitive processes by mainstream researchers (Ericsson & Simon, 1980, 1987; Leow & Morgan-Short, 2004).

Ericsson and Simon (1987) also discuss the possible types of observation that could be used for collecting specific types of data through verbalization. For instance, if the researcher's interest is in the answers generated by the subject and the total performance of the task, the process is called *performance observation*. However, if a researcher asks a subject to verbalize his thoughts as soon as they occur while he is solving the task, i.e., spontaneous verbalization, the researcher observes the processes and tracks them online. This type of observation is called *process observation*. In addition, the researcher can carry out *post-process observation*. Through post-process observation the researcher might obtain information about "memory for thought processes during the task, memory for presented information, and recollection of strategies used" (Ericsson & Simon, 1987, p.30). Nevertheless, post-process observation is vulnerable to the risk of change in the subject's mental processes, which may lead to the collection of incorrect performance data (ibid) and the emergence of additional information (Leow & Morgan-Short, 2004).

When collecting think-aloud data, other factors may define the quality of the data such as, instructions, warm–ups, and reminders. In giving instructions, for instance, phrases like "verbalize your thoughts" or "try to think aloud" are implicit requests to the subject to reveal his inner speech. However, if the researcher gives the instruction "Tell me whatever goes through your mind", it may capture thoughts that are both processed and waiting to be processed (Saravia, 1995, p. 30; Ericsson & Simon, 1987).

During a think-aloud session, a researcher may begin with a warm-up to acquaint research subjects with the process and instruments to be used, such as recorders, microphones, cameras, tasks, etc. (Ericsson & Simon, 1986). The researcher might ask subjects to perform a task that is either irrelevant, like solving an easy mathematical equation, or relevant to the main task in the study (Saravia, 1995). By warm-ups, the researcher can identify the factors that may either mislead the subjects or divert the think-aloud procedure from the desired path to unintended goals during the completion of the main task.

Finally, reminders are important. To guard against the risk of research subjects remaining silent, the researcher must always be present in the research milieu where think-aloud data are being collected. He can use reminders, such as "keep talking" or "what are you thinking about?" to encourage them to continue talking (Saravia, 1995, p. 30). The researcher should be cautious in selecting reminders since they may also alter subjects' verbalization. For example, unlike "Keep talking", which might be less influential on the processes inside subjects' minds; "What are you thinking about?" may produce explanations or descriptions (Saravia, 1995).

Overall, when collecting think-aloud data, a researcher should be careful about both the data types and protocols used to conduct the study since, as the previous findings and recommendations show, even choosing the wrong words, for example in issuing instructions or reminders, could result in invaluable or imperfect data.

2.7.1.3 Verbal Reports in Second Language Research

During the cognitivist era, second language researchers, like many others, were interested in exploring the learners' cognition and other procedures that were not easily observable and elicitable through the usual empirical research methods. The idea of exploring L2 learners' cognition by utilizing verbal report techniques is discussed in the immersive literature on SLA (Cohen, 1987; Katalin, 2000; Leow & Morgan-Short, 2004). In general, "verbal reports in SLA studies have been mainly used to explore the cognitive strategies adult learners use while reading an L2 text and their potential effects on subsequent comprehension" (Leow & Morgan-Short, 2004, p. 37). Verbal reports in L2 research were examined for the purposes of developing taxonomies of reading strategies, studying strategy transfer from L1 to L2 reading, defining the characteristics of both "good" and "poor" readers, understanding the effects of L2 learners' background knowledge on their reading comprehension, and exploring the cognitive processes of readers in L2 reading tests (Katalin, 2000). The think-aloud data obtained were unique for second language researchers. By referring to the data, second language researchers were able to present empirical evidence in support of their ideas and hypotheses (Leow & Morgan-Short, 2004).

Andrew Cohen (1987), in a study on the role of verbalization in second language research, identified three types of verbal reports: *self-report, self-observation,* and *self-revelation*. He described self-report as "learners' descriptions of what they do, characterized by generalized statements about learning behavior", self-observation as "the inspection of specific language behavior, […] while the information is still in short-term memory" (p. 84) either introspectively or retrospectively, and self-revelation as "a learner's report that is neither a description of general behavior nor based on the inspection of specific behaviors Rather it consists of "think-aloud" stream of consciousness disclosure of thought processes while the information is being attended to". He continues by introducing the factors that can affect the quality of verbal report data. These factors are: "the number of participants, the research context, the recency of the event, the mode of elicitation and response [either written or oral], the formality of elicitation [in a formal context, like a school, or out of the school], and the degree of external intervention [types of instructions and their roles in shaping the reports]" (Cohen, 1987, pp. 86-88).

To prevent faulty reporting and poor verbal reports, which Ericsson and Simon (1987) warn against, Cohen's categorization is recommended when the verbal report is the researcher's main data collection method. Thus, second language researchers who want to collect verbal report data are advised to optimize their research design with respect to Cohen's categorization.

	DESCRIPTORS											
	# Participants			Context		Mode						
		Indiv. +	Indiv				Elic.		Resp.			Degree of External
	Gp	Invest.	Alone	Dur. Class	Other	Recency	O	W	O	W	Formality	Interven.
TYPE OF DATA												
Self-Report	X	X	X		X	LO-HI	X	X	X	X	LO-HI	LO-HI
Self-Observation:												
introspection	X	X	X	X	X	HI	X	X	X	X	LO	LO-MED
retrospection	X	X	X	X	X	LO-MED	X	X	X	X	LO-MED	LO-MED
Self-Revealment:												
think-aloud		X			X	HI	X		X		LO	LO

KEY:
Participants: group, individual + investigator, individual alone
Context: during class, other
Mode: elicitation--oral, written; response--oral, written

FIGURE 4 Types of verbal report data in L2 and influential factors (Adopted from, Cohen, 1987, p. 85)

Overall, the think-aloud methodology can be considered beneficial for second language learning research. It provides valuable, in-depth insights for improving learners' attention to language input, assists them in reading and writing comprehensively, and supports them in learning to speak fluently and acquire new vocabulary items effectively. Moreover, the more second language researchers learn about learners' cognitive processes, the better their possibility of effectively adjusting their teaching practices (Cohen, 1987).

2.7.1.4 Advantages, Controversies, and Justifications

Like many other methods, verbalization and verbal reports have their disadvantages. The verbal report background theory is based on a distinction between working memory and long-term memory. Working memory has a limited capacity and heeded information is stored there first. For this reason, thoughts can disappear very quickly; therefore, only the reports that are verbalized very rapidly can be considered accurate in reflecting conscious thoughts. In other words, the researcher is limited to immediate verbal reports; otherwise, the accuracy of the data may be jeopardized. Moreover, many thoughts pass through the working memory simultaneously. Some of these thoughts may not be verbalized because either they are automatic or they are not yet processed (Charters, 2003). Thus, verbal reports cannot provide a clear

account of all the online processes happening in research subjects' minds during their engagement with tasks.

There is also the "homunculus" problem with verbal reports and introspection. In theory, introspection processes are separate and are not part of consciousness. If that is true, how can the research subject access these processes? In addition, "is the introspection process itself subject to introspection" (Someren, Barnard, & Sandberg, 1994, p. 30).

In second language learning research, the think-aloud protocol is a controversial issue. The debate centers around such topics as production as an object of study, the role of think aloud in L2 research, and the issue of reactivity (Bowles, 2010). Bowles discusses whether what study subjects produce as verbal reports is a proper object of study, and what the role of introspection actually is in L2 research; is it a complementary one? After discussing such questions, Bowles (2010) points out that it would be risky to make inferences based solely on verbal productions. Criticism is also leveled at think aloud, especially at introspective metalinguistic data, due to the issue of reactivity (ibid.). That is, "by thinking aloud, participants' internal processes may differ from what they would have been if they had not performed the verbalization" (Leow & Morgan-Short, 2004, 38). Thus, the issue of reactivity is always a consideration for the second language learning researcher when collecting verbal report data.

Despite the controversy over the think-aloud method, it can also be advantageous for language researchers (Bowles, 2010); for example, through verbal reports and think-aloud data we can access the unseen part of subjects' minds during their performance of linguistic tasks such as reading, writing etc., We can also learn more about individual differences by understanding to what extent background knowledge is activated while a person is engaged in, for instance, a reading task (Katalin, 2000), and finally, "the think aloud protocol yields detailed descriptions of task-induced […] behaviors and complexity in [subjects'] thoughts and that it also permits the effect of affective states on [subject]-task interaction" (Afflerback 2000, as cited in Yoshida, 2008, p. 200).

In this study, given the benefits of using verbal reports and the rich information they provide, I applied the think-aloud or - as it might also be termed - self-revelation protocol (Cohen, 1987). To answer the main question of this study as comprehensively as possible, I monitored and obtained the information heeded by the participants during task performance. Thus, to get the most out of my research, I collected concurrent non-metalinguistic think-aloud l data. I also designed and optimized the data collection method in accordance with mainstream researchers' recommendations (Erricson, & Simon, 1987; Cohen, 1987; Leow & Morgan-Short, 2004). Moreover, to enrich the elicited concurrent think-aloud data, I also conducted an "exit interview" (Charters, 2003), as recommended by mainstream think-aloud researchers. There are good reasons for conducting an exit interview. Think-aloud data is only a partial description, since thoughts or processes do not remain in working memory long enough to be verbalized; thus, some verbalization may not be performed or it may be performed but forgotten (Leow & Morgan-Short, 2004). Moreover, verbalized

data may differ in both quality and quantity owing to various factors such as individual differences, the protocols used etc. (Ericsson & Simon, 1980). Hence, to attain a deeper understanding, an exit interview is recommended (Charters, 2003; Ericsson & Simon, 1980). Although, through the questions asked, the exit interview may be biased towards the researchers' intentions, the retrospective data obtained will be valuable when combined with the concurrent think-aloud data (Nunan, 1992 as cited in Charters, 2003). In this study, the exit interview was conducted to add depth to the concurrent think-aloud data and to audit the validity of my interpretations of the concurrent data.

Consequently, by applying a think-aloud protocol, I can not only report what level of involvement turned out to be the most beneficial for acquiring the target vocabulary items but also describe why that level is optimal. In addition, the approach provides an opportunity to compare involvement load levels in their effects and to identify their similarities and differences.

2.7.2 Exit-interview

"The interview is the most often used method in qualitative enquiries" (Dörnyei, 2007, p.134). Interviews can be classified by type and structure. For example, the four main types of interviews are the single- or multiple-session interview, structured interview, semi-structured interview, and unstructured interview (Dörnyei, 2007). Researchers select the type of interview best suited to the purpose of the data collection. For example, in a single- or multiple-session interview, the researcher may obtain enough data in the first session or may conduct further sessions depending on the depth and breadth required of the data. In the second type, degree of structure, a highly structured interview follows pre-determined guidelines and questions are selected based on pre-identified criteria. Unlike a structured interview, an unstructured interview "allows maximum flexibility to follow the interviewee in unpredictable directions, with only minimal interference from research agenda" (Dörnyei, 2007, p.135). It is believed that an unstructured interview may be comfortable for the interviewee and lead to the elicitation of deeper and broader data. And finally, between the two extremes, i.e., the structured and unstructured interview, is the semi-structured interview, which is the commonest type in applied linguistics. In a semi-structured interview, "although there is a set of pre-prepared guiding questions and prompts, the format is open-ended, and the interviewee is encouraged to elaborate on the issues raised in an exploratory manner" (Dörnyei, 2007, 136). In other words, it not only guides the interviewee, based, for instance, on scientific agendas for eliciting critical and valuable data, but it also gives the interviewer enough freedom to express new ideas during the interview. Therefore, the type of interview depends on the manner and purpose of the data collection.

In this study, two types of questions were asked in a researcher-designed semi-structured interview. The first type were general questions aimed at eliciting more information about the processes and strategies that participants used for either learning or understanding the target new words during their task

performance such as, "Has this video game helped you to remember / learn new words easily? If so, how?". The second type were questions formulated on the basis of the *Cognitive Theory of Multimedia Learning* (Mayer & Moreno, 2003). Thus, I adjusted my questions to its underlying assumptions, which are determining factors in either the success or failure of any multimedia learning context. These questions not only further informed me about the participants' online processes during their task performance but also assisted me in checking and controlling for the effect of the multimedia learning factors. In this theory, the human mind works based on three assumptions: *dual channel*, *limited capacity*, and *active processing*. Mayer and Moreno (2003) explain these as follows: 1) dual channel means that "humans possess separate information processing channels for verbal and visual material"; 2) limited capacity refers to the fact that "there is only a limited amount of processing capacity available in the verbal and visual channels"; and 3) active processing assumes that "[meaningful] learning requires substantial cognitive processing in the verbal and visual channels" (p. 44). These assumptions are drawn from different theories, for example, the dual channel assumption is based on Paivio's dual coding theory, which was explained in section 2.4.3. The idea of limited capacity is adopted from Sweller's (2010) cognitive load theory, where he posits that either too high or too low a cognitive load may have negative effects on learning comparing to balanced cognitive load, which he calls a germane load. Mayer and Moreno (2003, p. 44) associate active processing assumption with "Wittrock's (1989) generative-learning theory and Mayer's (1999, 2002) selecting–organizing–integrating theory of active learning". The processes, in this assumption, "include paying attention to the presented material, mentally organizing the presented material into a coherent structure, and integrating the presented material with existing knowledge". According to Mayer and Moreno's (2003) theory, both the visual and the auditory inputs can be modified and controlled, in all three assumptions, either to increase the applicability and effectiveness of the multimedia materials used or to reduce the effect of undesired factors that decrease the quality of learning via multimedia in educational contexts. Bearing all the above in my mind, I formulated my interview questions.

In sum, I designed a semi-structured interview to obtain more and deeper information about the processes and strategies that participants follow in completing set vocabulary acquisition tasks.

3 DIGITAL GAMES, VOCABULARY, AND ACQUISITION

Digital games are currently a highly popular form of media entertainment. From the 1980s to the end of the first decade of the new millennium, the average amount of time spent playing digital games has risen from 4.5 to 9 hours per week in the United States (Dehaan, Reed, & Kuwada, 2010). The growing global popularity of digital games among people has stimulated the curiosity and interest of researchers, who see digital games as another avenue for research and study in a safe 'virtual environment' (Kirriemuir, 2002).

Following this new trend, applied linguists and language acquisition experts have also researched the digital game phenomenon. They have found that digital games can offer precious opportunities for language acquisition via such factors as repetition, contextual clues, transfer, motivation, awareness, controllability, active engagement, comprehensible target language input, learner-centered situations, negotiation, reduction of affective variables, collaborative dialogues, community of practice, experimental learning, mediation, and motivation (deHaan, 2005; Yildiz & Trugut, 2009; Pasfield-Neofitou, 2014).

Among the opportunities that digital games offer language learners is that of vocabulary enhancement, a crucial component of language acquisition. However, owing to the multidimensional and complex nature of vocabulary acquisition, the digital game effect may not be as straightforward as one might be led to expect. However, before addressing the reasons why, in the realm of vocabulary acquisition, digital games may have differing effects, I will discuss the nature of digital games, their roles in acquisition, and the rationale behind them in education, general learning, language acquisition generally, and vocabulary acquisition. Thus, in this section of the thesis, I provide an overview of digital game studies, first discussing the terminology used and then reviewing the literature.

3.1 What is a game?

First, I need to clarify what constitutes a game in my study. According to Arjoranta (2014), this is not an easy task, as, despite the multiple definitions available, no consensus has as yet been reached. Wu, Franken, and Witten (2012) define a game, by and large, as an activity that not only entertains and engages but also offers different types of challenges. Caillois (1961 as cited in Garris, Ahlers, and Driskell, 2002) defines a game as an activity that has the following attributes: enjoyable, voluntary, dominated by its own rules, unpredictable, and fruitless, meaning that it does not have any useful outcome. Garris, Ahlers, and Driskell (2002), expanding Caillois' definition, add the following six dimensions as the main attributions of a game: fantasy, rules/goals, sensory stimuli, challenge, mystery, and control. Furthermore, they add that "a game is a system that is a real world by its own right and does not simulate any other systems or worlds" (pp. 442- 443). Finally, they posit that the more of these aspects an activity incorporates, the more game-like it becomes. However, none of the above can be considered a comprehensive definition of a game according to Arjoranta (2014), who also points out that it is not necessary to reach a comprehensive and conclusive definition.

Arjoranta (2014) proposes an approach that he calls Wittgenstein approach for overcoming the difficulty of game definition instead. He explains that, the games must be defined according to their resemblances and features that they share with context and cultures that they are applied to. After spotting the common aspects between the games and cultural contexts, the points of emphasis will be specified. In that case, defining a game can be logically acceptable and applicable to that context. In this study, the digital version of games is focused on. Thus, I try to define digital games as the main form of game in this study.

Digital games are computerized and digital forms of entertaining activities (Prensky, 2007). Context-wise, for this study, a digital game is discussed in a learning context. Thus, based on Arjoranta's suggestion, a proposed definition needs to incorporate both the learning context and the digital game. Therefore, the digital game, in this study, can be defined according to Hainey, Connolly, Boyle, Wilson, & Razak (2016, p. 204) as an "innovative learning approach derived from the use of computer games that [lead to] educational [effect] or different kinds of software applications that use games for learning and education purposes such as learning support, teaching enhancement, assessment and evaluation of learners".

3.2 What is Digital Game-Based Learning (DGBL)?

The emergence of digital games and their impacts on people's lives has also led to the coining of new terminologies. In digital game studies, the term digital game-based learning, or DGBL, is ubiquitous. The purpose of DGBL is to use

digital games to integrate the end user and a specific aim, and it uses the elements or characteristics of digital games to create an engaging, fun, motivational, and suitable learning experience that will assist learners in becoming experts in a specific skill (Zin, Jaafar, & Yue, 2009). Prensky (2007) observes "that DGBL is stealth learning because it is a new paradigm that is gradually emerging" (p. 19) in order to elevate learning via play. He adds that it is a process aimed at removing pain from work and substituting it with enjoyment. The outcome of this process is getting better at something without undergoing hard labor but instead having fun. (Prensky, 2007).

Razak, Connolly, and Hainey identify two categories of DGBL: "learning through game playing" and "learning through game making" (2012, p. 35). They report that the most common type of DGBL is learning through playing. A more detailed categorization of DGBL is posited by Ermi, Heliö, and Mäyrä (2004) who see DGBL as somewhere between digital games and education programs. They believe that the best description of DGBL is edugames, or games designed for educational purposes. They add that an edugame is different from edutainment in how the content is taught. In edugames, a game is designed for teaching specific content, whereas in edutainment, the elements and features of digital games are added to an educational activity to make learning more game-like and fun.

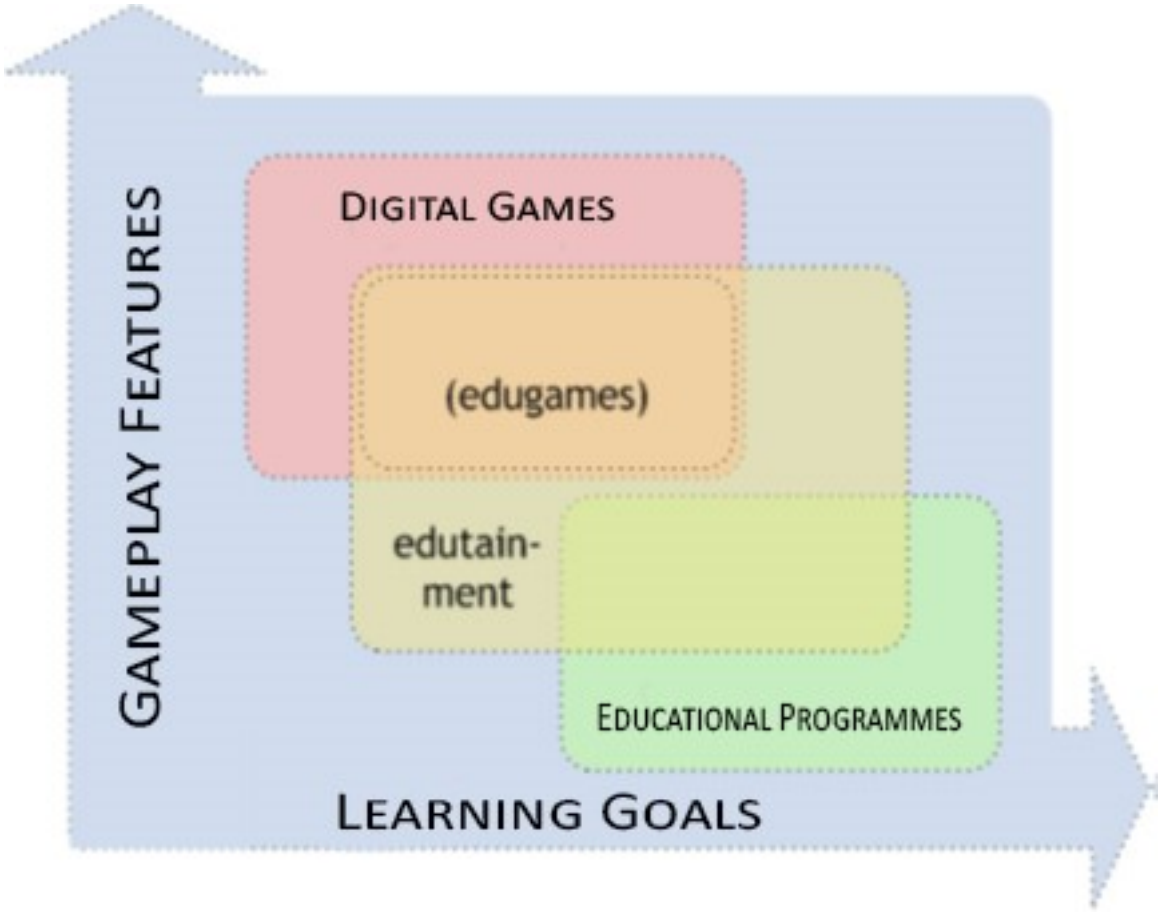


FIGURE 5 Categories of DGBL (Adapted from Ermi, Heliö, & Mäyrä, 2004, p. 62)

In any kind of DGBL, the following benefits can be expected: the hardest and dullest subjects become easier to learn via various motivating elements; the creation of digital games make learning fun for all groups of learners and teachers; and each learner has his or her own trainer who is the best at teaching a specific subject (Prensky, 2007). The effects of these outcomes increase the quality of learning since they add motivation, entertainment, involvement, engagement, and induce a state of flow (Kerrimiur, 2002; Garris, Ahlers, & Driskell, 2002; Gee, 2003; Bowman, 1982, as cited in Squire, 2003; Jackson, Dempsey, & McNamara, 2012).

"Digital games are viewed as a tempting form of student engagement, given the power of games to profoundly immerse and engage players" (Shahriarpour, 2014, p. 1739). DGBL opens new doors and offers new opportunities for every learner. For example, it may help preschoolers to learn the alphabet and reading skills. It helps in learning the curriculum: elementary learners were helped to score higher than others on the K-6 curriculum simply by playing PlayStation (Sony entertainment gaming system). A digital game can also assist students to learn typing or, say, electoral politics. *SimCity* has helped traders to increase their financial skills, and policy makers to understand the health care system. DGBL has assisted engineers to learn CAD and to improve their skills. By integration of a war simulator, military personnel can learn how to fight and how to survive under various physical and psychological conditions in a subsequent real conflict (Prensky, 2007).

> [In DGBL] learners are encouraged to combine knowledge from different areas to choose a solution or to make a decision at a certain point, learners can test how the outcome of the game changes based on their decisions and actions, learners are encouraged to contact other team members and discuss and negotiate subsequent steps, thus improving, among other things, their social skills. (Pivec, Dziabenko, & Schinnerl, 2003, p. 217)

In other words, self-efficacy, interest, engagement, decision making, and self-regulation can be invigorated in DGBL contexts. Moreover, DGBL implements elements such as game play, challenge, fun, enjoyment, feedback, incentives, systematic task difficulty, control, and the role of environment in order to support the indicated individual attributes either to emerge or to replenish (Rankin, Gold, & Gooch, 2006a; Jackson, Dempsey, & McNamara, 2012).

The possible benefits of DGBL have also been studied empirically. To examine the effect of DGBL, researchers have mainly used digital games, both edugames and edutainment, for delivering and teaching content. In these studies, research participants have been asked to play digital games in order to learn new content. Overall, DGBL has been found to have beneficial effects.

For example, Squire (2005), in his comparative case study, investigated the effect of a historical simulation digital computer game, *Civilization III*, for teaching history and geography in two schools. The 33 participants were school students in history classes. Through interview, field notes, and observation, he found that the students retained and understood the content of history courses better after playing the digital game. They also more effectively learned details

about the geography and chronology of historical events. He concluded that incorporating digital games into the curriculum promotes better understanding of content in history courses.

Schlickum, Hedman, Enochsson, Kjellin, and Felländer-Tsai (2009) tested the effect of digital games making high visual-spatial and cognitive demands on the endoscopic surgical simulator performance of thirty surgical novices. They hypothesized that the participants with experience of playing digital games would outperform non-players in a virtual reality surgical simulation exercise. The 30 novices in the experimental group played *Half-Life*, a first-person shooter game with high visual-spatial demands, and Chessmaster, a chess simulation game with high cognitive demands, for 10 sessions. The control group, containing 10 surgical novices, practiced surgery through the virtual reality surgical simulator without playing any digital games before the simulator exercise. Their results showed that the digital game group outperformed controls in the surgical simulation exercise. This result was attributed to improvements in both the visual-spatial imagination and cognitive skills of the game group participants.

In a study by Johnson (2007), 20 US marines played *Tactical Iraqi*, a digital game designed by the *Tactical Language and Culture Training System* for helping soldiers to acquire communicative skills in Iraqi Arabic. Only one of the marines, who had been deployed in Iraq for a while, had any experience or familiarity with the target language. After 50 hours of training, their language level was tested using evaluative software. The results showed a mean score of 3.73 out of 5, which was an indicator of progress in the acquisition of communicative skills in Iraqi Arabic, both in cultural awareness and linguistic knowledge. Moreover, in line with the test results, 78% of the participants felt they had gained considerable communicative skills in Iraqi Arabic.

Thus, it can be inferred that the use of appropriate digital games in learning situations can significantly promote learning. DGBL can also engage learners in both critical and active thinking (Rankin, Gold, & Gooch, 2006b), leading to meaningful learning, which is one of the desired outcomes of any learning-oriented program.

3.3 How do digital games enhance learning?

The concepts underlying DGBL were introduced in the previous sections. In this study, digital game-related terminology was considered important in assisting understanding of the potential role of digital games in learning, especially in elevating learning. In this section, I discuss how digital games can assist or improve learning in general. The purpose of this section is to provide a foundation for better understanding the role of digital games in language learning and, ultimately, vocabulary learning, the primary focus of this study.

Broadly speaking, there are two different views on how digital games support learning. Some researchers argue that digital games are beneficial for different aspects of learning due to their internal elements such interactivity, challenge, visuals etc. (Gee, 2003; Pivec, Dziabenko, & Schinnerl, 2003; Squire, 2003; Shaffer, Squire, Halverson, & Gee, 2005; Johnson, Vilhjálmsson, & Marsella, 2005; Dondlinger, 2007; Figg & Jaipal, 2009; De Freitas & Maharag, 2011). They posit that the positive cognitive and motivational effect of internal elements can improve learning. Other researchers, such as Gee (2003), and Becker (2008), consider the internal design of digital games to be the main factor in boosting learning. They believe that the internal elements of digital games interact with each other in a way that leads to the emergence of a system that largely conforms to current major theories of learning such as constructivism.

With respect to the first view, Squire (2008, p. 15) speculates that "the emergent paradigm of game-based learning is built on a number of principles" and elements the presence of which in digital games facilitates successful and efficient learning. Elements such as interactivity, rules (such as in game physics), goals (such as defeating an evil boss), challenge (such as how to unlock a jigsaw box), risk (such as jumping over a wide pit), fantasy (such as gaining magical powers), curiosity (such as exploring an abandoned cave), control (the ability to control one's actions), gameplay (or how the game can be played), different types of feedback, such as evaluative, interpretive, supportive, probing (De Freitas & Maharg, 2011), affordances, fun, replay motivation (such as the feeling that I have to beat my enemy after lots of attempts), difficulty and complexity, and failure and choice (or trial and error), are among the most often adduced (Pivec , Dziabenko, & Schinnerl, 2003; Johnson, Vilhjálmsson, & Marsella, 2005; Squire, 2005). However, perceptions on how those elements assist learning differ.

For Hamari, Shernoff, Rowe, Coller, Asbell-Clarke & Edwards (2016), digital game elements create a virtual world system. The virtual world of a game, which is a systematic combination of such elements, is praised as the major component that turns digital games into potential contexts for learning. It is posited that virtual worlds speed up learning by developing situated learning, experimenting, experiencing, forming new and powerful identities, and integrating knowing and doing that leads to situated understanding (Pivec, Dziabenko, & Schinnerl, 2003; Johnson, Vilhjálmsson & Marsella, 2005; Shaffer, Squire, Halverson, & Gee, 2005). Hamari et. al. (2016) add that when learning is situated, the possibility of hypothesizing, probing, and reflecting upon the virtual world of digital games increases. Thus, learners have the opportunity to modify various parameters, to obtain new views on the phenomena, to monitor how a system behaves over time, to analyze and criticize a system, and to acquire a visual picture of a system and many other features that assist the learners' learning approaches and outcomes (Squire, 2003, p. 55).

In a second speculation on how internal elements assist learning, Alexiou and Schippers (2018) propose that the internal elements of digital games support learners' cognitive-motivational engagement, which is an important factor in learning. They continue as follows:

> The cognitive and emotional engagement of players are a function of game elements that are nested within three main layers: the game system (rules, mechanics), narrative (theme, story, characters) and aesthetics (Audiovisual elements, fidelity, aesthetic choices). When it comes to learning applications, cognitive engagement refers primarily to the focus of attention, while emotional engagement stresses the role of emotions and feelings in supporting the desired cognitive processes. (Alexiou & Schippers, 2018, p. 2550)

This idea supports the conceptual framework of gamification and learning proposed by Hamari, Koivisto, and Sarsa (2014), who also see game and gamification-assisted learning as comprising three phases. In the first phase, games provide motivational resources and affordances that, in the second phase, can cause psychological changes in learners. In the third phase, these psychological changes lead to behavioral changes and outcomes (Hamari, Koivisto, & Sarsa, 2014). These behavioral changes, if guided, are considered as learning. Therefore, it can be inferred that the internal elements of digital games enhance learning by enhancing motivation. This in turn not only results in emotional engagement but also supports cognitive engagement, or cognitive processes. These processes can be intentionally invoked to produce a desired behavior, i.e. learning.

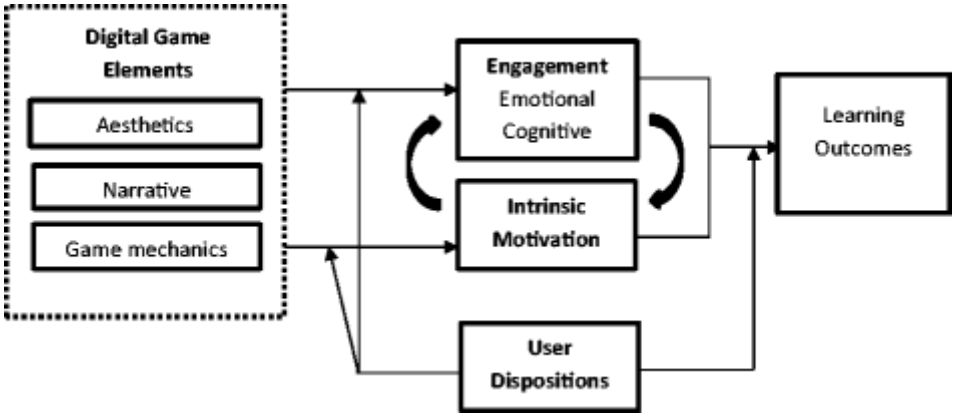


FIGURE 6 Game elements and learning. (Adapted from Alexiou and Schippers, 2018, p. 2547)

A third speculation is about the role of internal elements, specifically challenge, in boosting learning by manipulating the player's motivational levels through enjoyment and emotions. Jackson, Dempsey, and McNamara (2012) discuss entertainment in digital games as a source of joy. Moreover, the existence of elements like fantasy, rules/goals, sensory stimuli, challenge, mystery, and control generate entertainment (Garris, Ahlers, & Driskell, 2002). Setting the internal element of challenge at an appropriate level was found to generate the emotional state of flow (Ebrahimzadeh & Alavi, 2016). The state of flow is defined by eight components of enjoyment.

> First, there should be a chance to complete the task or challenge. Second, it should be possible to concentrate. Third and fourth, concentration is usually realized through having clear goals and receiving immediate feedback. Fifth, task involvement is so deep, yet effortless, that removes daily worries and frustrations. Sixth, the enjoyable

> experience enables people to feel some sense of control over their activity. Seventh, the sense of self is forgotten during flow and emerges stronger afterwards. Finally, time distortion happens with hours passing by in minutes or vice versa. Should all these elements be present, their combination causes a sense of deep enjoyment that is so rewarding people feel that expending a great deal of energy is worthwhile simply to be able to feel it (Czikszentmihalyi, 1991, p. 49)

Hamari et. al. (2016) studied the state of flow and the internal elements of digital games. They found that, among many other internal elements of digital games, the element of challenge was the major contributing factor in generating flow in learners. Their statistical analyses also revealed that, in digital game-based learning, the element of challenge had a direct effect on learning and mediated the effect of engagement. Furthermore, they found that the learners who experienced flow during their game play reported a higher level of emotional enjoyment. During the state of flow, learners were also more engaged and motivated. To summarize, digital games assist learning because they provide enough challenge for learners to experience the state of flow. When they experience flow, learners feel more emotionally engaged and motivated. According to Alexiou and Schippers (2018), a high level of motivation supports cognitive engagement or processes. Thus, learning outcomes are effective and noticeable.

A fourth speculation on the role of the internal elements of digital games concerns how combinations of internal elements, their interactions, and their situated relevance can influence learning. Ang and Zaphiris (2006) posit that digital games may influence learning because the components of the digital games can interact in a manner that paves the way for internal elements to affect learning via the enhancement of motivation. They make this claim by reference to the ideas of narratology and ludology. They defined ludology in computer games as the study of activity, and narratology in computer games as the study of stories. Drawing on the concept of entertainment and enjoyment (Jackson, Dempsey, & McNamara, 2012), the ludologists saw gameplay, or any activity allowed by the rules of the game directed at achieving a goal in the game's universe, as the main source of entertainment and enjoyment, while narratologists identified the story or narrative, or series of events, as the main source of entertainment in digital games (Ang & Zaphiris, 2006, pp. 2-3). They supported their view by reference to the components of gameplay. They argued that rules were units that formed the basic structure of any gameplay. Knowledge of these rules enabled the gamer to play and the game and to be playable. Rules were categorized into two types. They were either *paidea* rules "that describe the semantic of the game" or *ludus* rules "that describe the structure of the game" (Ang & Zaphiris, 2006, p. 6). The paidea rules, such as how far Mario (in the *Super Mario Bros* series) can jump, are made by the game designers. The ludus rules are flexible ones; although stipulated by the game designers, they can be violated by the player. For example, while the ludus rule of Mario is to get to every level by overcoming a pre- identified number of obstacles, the player can violate this rule by playing around in some section of a level. However, every digital game has a narrative of some kind. To allow for both gameplay and narratives, Ang and

Zaphiris (2006) relate the paidea rules to the narrative setting and the ludus rules to the narrative plot. On this view, the paidea activities were determined by "environment and action" (ludus rules) (Ang & Zaphiris, 2006, p. 7). Moreover, spatiality, or "the space of the narrative", and fabula, "or the actions that might happen in these spaces", were considered as two main components of narratives (Ang & Zaphiris, 2006, p. 7). Ang and Zaphiris (2006, p. 19) hold that the interaction between the components of gameplay and narrative in digital games can lead to a high quality learning because "knowledge is constructed instead of being transmitted [, … this combination makes learning] it is also motivating, where it challenges the learners, intrigues their curiosity and brings about fantasy". Based on this final speculation, it can be proposed that the internal components as well as elements of digital games can assist learning through entertainment, enjoyment and motivation. Thus, overall, it can be concluded that, from every perspective on the internal elements, digital games influence learning by directly manipulating the level of motivation and indirectly by supporting cognitive factors via motivation.

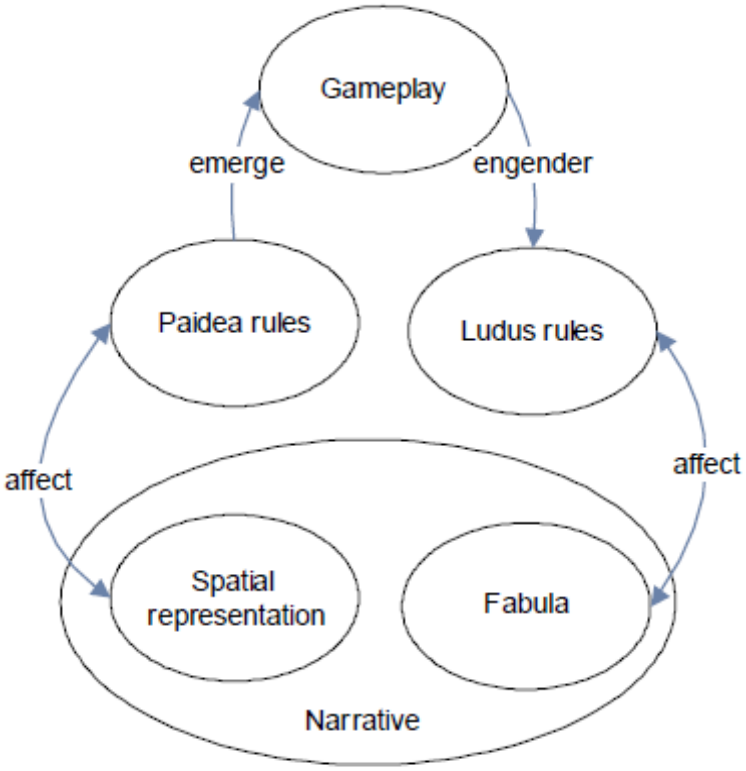


FIGURE 7 How the elements of gameplay and narrative interact (Adapted from Ang & Zaphiris, 2006, p.8)

As discussed above, there is a second perspective on how digital games might enhance learning. The first viewpoint argues for the efficacy of digital games in learning based on the positive impact of individual internal elements rather than these as a whole or systematically. However, in the second point of

view, the success of a digital game in enhancing learning is due to its internal design, that is, the effect of its internal elements as a whole; this view is informed by such contemporary learning theories as constructivism. In other words, from the first point of view, the success of a digital game is evaluated in accordance with the quantity of its internal elements. For example, digital game A is more effective for learning than digital game, B because it contains, for example, an additional two effective internal elements. However, from the second point of view, the effect of the existing internal elements is considered and discussed as a whole. For example, digital game B might have fewer internal elements but the effect that they may have as whole on learning is greater than the sum of the individual effects of digital game A.

Unlike Gee (2003), who proposed, based on 36 principles extracted from previous learning theories, that digital games might assist learning literacy, Becker (2008) posits that digital games enhance learning because they follow the principles, events, and components of six theories of learning, viz., Gagné's nine events of instruction, Reigeluth's elaboration theory, Merrill's first principles of instruction, Constructivist learning environments, activity theory, and problem-based theory. Thus, the role of digital games in assisting learning could be studied through the lenses of those theories of learning.

Generally, digital game-based learning has been considered more from a constructivist viewpoint more than any other because digital games supply learners with an environment in which they can experience "playful imagination -allowing learners to simulate and experiment with real-life scenarios-, social interaction with more capable peers - allowing learners to enhance their cognitive understanding-, and intrinsic motivation - required for sustaining involvement in self-regulating learning over time" (Alexiou & Schippers, 2018, p. 2549). In such a sandbox environment, "players are allowed to experiment and construct meaning out of their cognitive and emotional experiences" (Alexiou & Schippers, 2018, p. 2547).

To clarify the association between the internal design of digital games and constructivist view of learning, Becker (2008) looks at the gameplay of the *Super Mario Bros* game series. She explains that an inactive gamer was rarely found in every Mario game observed (Active principle: learning should be engaging and should keep learners active). By passing through each level, the gamer develops either a new or an old skill (Constructive principle: new ideas and prior knowledge should be integrated for making meaning). The community of Mario fans aid each other with suggestions on websites and in articles and, magazines etc., and they share their skills with each other in cases of a demanding problem arising at a specific stage of any Mario game (Collaboration principle: learners should work in communities to obtain support and an opportunity for collaborative knowledge building). This game has a goal and, by supplying elements of fun, it encourages gamers to continue playing (Intentional principle: if a learning task is goal-oriented and intentional, learning will be boosted). Mario games provide various types of challenge for the gamer (Complex principle: both complex and simple problems, in learning tasks, can be engaging and can

enhance learning). In the *Paper Mario* series, Mario visits different people and uses his repertoire of extraordinary abilities to save princess Peach. To gain those skills and abilities, the structures and components of the tasks and levels are designed to assist the gamer reach the maximum level of skill knowledge (Contextual principle: learning should take place in a meaningful world or task). Because learning is the outcome of interaction, we are returned to the Mario community where gamers interact to obtain knowledge and improve their skills (Conversational principle: the process of learning should be dialogical and conversational). The outcome of the interaction among Mario fans is the acquisition of further knowledge of what is right and wrong in the game. Some comments will be criticized by other fans because they know a better way to reach a specific goal in the game (Reflective principle: learners should know what they do in a task, what processes they have followed, what strategies have been useful, and what goals they have achieved).

Thus, according to second point of view, the main reason for the effectiveness of digital games in learning is that the games are internally designed so that their internal elements work systematically together. It is this feature that enables them to have a positive effect on learning and the construction and transfer of knowledge.

> In other words, [the internal design] grant[s] the players the ability to a) retrieve relevant knowledge from memory, b) determine the meaning of instructional messages, c) apply a procedure to a given situation, d) identify the constituent parts of the material, how they relate to each other and their role in the overall structure or purpose, e) make judgments based on criteria and standards and f) create something novel out of existing elements. (Alexiou & Schippers, 2018, p. 2548)

3.4 What is digital game-based language learning (DGBLL)?

Language acquisition experts have also studied the educational efficacy of digital games for language acquisition and learning. As mentioned earlier, new trends and technologies generate novel terminologies. One such term is *digital game-based language learning* (DGBLL), which refers to any educational use of digital games for the purpose of language learning and extends the earlier term DGBL. However, with respect to the nature of integration, other definitions of DGBLL exist. Ang and Zaphiris (2008) state that DGBLL includes two perspectives: 1) how to use digital games as virtual environments that facilitate language learning, and 2) how to use digital games as instruments for providing and enhancing opportunities for, e.g., collaborative learning. Reinhardt and Sykes (2012) consider that the definition by Ang and Zaphiris (2008) is the most comprehensive. They also see two categories of DGBLL: game-enhanced and game-based DGBLL and present a taxonomy of DGBLL research and practice. Any curriculum in which a digital game is used as a side tool for developing learners' language proficiency is a game-enhanced language learning curriculum whereas if the gamer learns the language by playing a game, it is a game-based

language learning curriculum. The categories are presented in detail in the following table:

TABLE 6 Guiding questions for game-mediated l2 learning program research and practice (Adapted from Reinhardt & Sykes, 2012, p. 33).

	L2 learning focus	**L2 pedagogy focus**
Game-enhanced: working with vernacular games, or commercial games	How does game-mediated L2 learning occur in the wild?	How can vernacular games be pedagogically mediated for L2 learning and teaching?
Game-based: working with educational and L2 learning purposed games (i.e., synthetic immersive environments)	How do specific game designs afford particular L2 learner behaviors	How can game-based environments be designed to incorporate and/or complement L2 pedagogical use?

By and large, digital game-based learning is deemed effective for language learning (Chian-Wen, 2014; Hung, Yang, Hwang, Chu, & Wang, 2018). Chian-Wen (2014), in a meta-analysis, found that both kids and adult language learners benefited from DGBLL. Referring to the large positive effect sizes that he found between the presence of digital games in language learning and learners' long-term engagement, he suggested that the integration of digital games in the language learning process can facilitate the acquisition of procedural knowledge.

deHaan (2005), in a case study, taught Japanese to a 27-year-old male American intermediate Japanese language learner with the help of a baseball digital game. After a month, using self-report, observation, interview, and listening and reading tests, he assessed the subject's possible acquisition of Japanese and concluded that by playing a baseball digital game, the learner had improved his listening, reading, and Kanji character recognition. deHaan (2005) suggested that the repetition, contextual language, and simultaneous aural and textual representation of language in the game were the factors that contributed to the subject's language learning and improvement.

Suh, Kim, and Kim (2010) utilized a MMORPG game to teach elementary-level English in all four skills, controlling for variables such as gender, prior knowledge, motivation, self-directed learning skills, computer skills, game skills, computer capacity, network capacity, and computer accessibility. They compared MMORPG-based English learning to face-to-face traditional English instruction. Participants were 220 fifth and sixth graders from 5 different schools in Seoul. 118 participants were assigned to the digital game group and 102 to the face-to-face group. The experimental group used the digital games twice a week for 2 months in order to complete the in-game tasks. The control group were taught using a course book and in-class language learning activities. The results revealed the superiority of the MMORPG-based English learners in listening, reading, and writing. Among the variables controlled for, they found prior knowledge, motivation, and network speed had a considerable influence on learning English via MMORPG.

Neville, Shelton, and McInnis (2009), in a mixed-methods study, taught German vocabulary, reading and culture via an interactive fiction game to university students. Participants were 15 third-semester German-language learners assigned to either an experimental group (7 participants), and a control group (8 participants). The experimental group received the story by play the interactional fiction game and the control group by traditional instruction. Both groups were presented with the same vocabulary items. Comparison of the groups' performance showed that the experimental group outperformed the controls. The questionnaire results also indicated that most of the participants preferred digital games to traditional instruction. Increased mental effort and both the immersive and interactive environmental features of the interactive fiction game were speculated as the possible factors that led to the better performance of the digital game group. Finally, they reported better acquisition of new vocabulary items by the digital games group.

Anderson, Reynolds, Yeh, and Huang (2008) evaluated the effect of digital games on language learners' listening comprehension by using the digital game *America's Army*. They recruited 29 undergraduate non-English speaking Taiwanese participants, both males (17) and females (12), for their study. Participants were randomly assigned to either the experimental group, who played the digital game, or control group, who received instruction in a traditional listening format. In a pre-test-treatment-post-test design, 10 vocabulary items were taught to both the experimental group and the treatment group. The vocabulary items were the same and they were used both in the tutorial section of the digital game and the control group task. Both groups had to listen to instructions and complete the assigned tasks. After assessing the groups' listening comprehension in both pre- and post-tests, they concluded that digital games can be improve listening comprehension. Students also reported positive feelings about language learning via digital games.

It may be asked: how is that digital games can positively affect language learning? To answer this question, the first viewpoint in digital games and general learning, i.e. the effect of internal elements individually, has been widely researched. That is, the effectiveness of digital games on language learning has been studied from a cognitive-motivational perspective focusing on the internal elements of digital games separately rather than holistically (Hung, Yang, Hwang, Chu, & Wang, 2018). Moreover, following the internal element perspective, digital games' internal elements are discussed as supporters of factors whose effectiveness in language learning, has been studied and confirmed. The features that digital games' internal elements can offer for language learning are e.g. repetition, contextual clues, transfer, motivation, awareness, controllability, active engagement, comprehensible target language input, learner-centered situations, negotiation, reduction of affective variables, collaborative dialogues, community of practice, experimental learning, mediation, willingness to communicate, immersion, and motivation (deHaan, 2005; Yildiz & Trugut, 2009; Pasfield-Neofitou, 2014; Hung, Yang, Hwang, Chu & Wang, 2018). Peterson (2010), emphasize that various advantages of digital

games for language learning may be due to different internal features of digital games. He hypothesizes that these features may promote language learning by supporting the factors that play crucial roles in the process of language learning. He lists them in the following table (Table.7).

TABLE 7 Hypothesized advantages of digital games in language learning (Adapted from Peterson, 2010, p. 432)

Design feature	Hypothesized advantages
Network-based real-time text and voice chat	Access to diverse groups of interlocutors, including native speakers Multiple communication channels provide real-time feedback Exposure to the TL The presence of text and scrolling supports monitoring Extensive opportunities for purposeful TL use and reuse in an authentic and engaging communicative context Practice in the four skills Opportunities to engage in co-construction, negotiation, and the development of communicative competence Learner-centered interaction encourages active participation Enhanced cross-cultural knowledge
Challenging theme and goal-based instruction	Motivation enhanced Enjoyment Situated learning Community formation Development of collaborative social relationships
Personal avatars	Enhanced immersion Opportunities for role-play and risk-taking Reduced inhibition and social context cues

Pasfield-Neofitou (2014), in her meta-analysis, concluded that the presence in digital games of both psychological and sociocultural factors that could play a major role in language learning boost the efficiency of digital games in language learning contexts.

However, Reinhardt and Thorne (2016, as cited in Sundqvist, 2019), pointed out that language learning is not the sole outcome of the internal elements of digital games. They argue that language is learned not from digital games only, but also from the discourse surrounding digital games. In this regard, Scholz (2017) recently offered a new perspective on how digital games may enhance language learning. His view is much closer to second perspective on digital games and general learning, according to which the relationship can be explained by learning theories or by studying the digital game effect as a whole. Scholz (2017) discusses how out-of-school digital game play facilitates and develops language learning can be explained in terms of the Complex Adaptive Systems framework (CAS). He states that

> complex adaptive systems view language both cognitively and socially, allowing for detailed, non-reductionist analyses that take into account as many factors and

> variables as possible to understand the change and SLD [second language development] that occurs, thus resulting in a unified approach of the system as a whole, as opposed to a singular variable or aspect of the system. (Scholz, 2017, p. 42)

To test this idea, he recruited 14 L2 learners of German (12 males and 2 females). They were asked to play *World of Warcraft* for 202 hours out of class in German. Moreover, they were asked to discuss *World of Warcraft*-related topics. The data were elicited by background information questionnaires, in-game and in-person communication logs, and interviews. Analysis of the qualitatively collected data showed that the simultaneous existence of three factors generated a condition that facilitated SLD in DGBLL in the long run. Those factors were gameplay, communication, and iteration. He concluded that, from a CAS perspective, the existence and interaction of those factors generates a system that supports sociocognitive processes that in turn facilitate linguistic construction.

3.5 Digital games and vocabulary acquisition

Much digital game and language learning research has been devoted to experimental studies on the effect of digital games on vocabulary acquisition. Meta-analyses (Chiu, Kao, & Reynolds, 2012; Chian-Wen, 2014), have shown that hypotheses on the positive effects of DGBLL in vocabulary acquisition have been empirically supported. For example, the results of the meta-analysis by Chen, Tseng, and Hsiao (2018) showed that DGBLL improves vocabulary acquisition.

Thus, the results of previous studies mostly report positive effects of DGBLL on vocabulary acquisition. For instance, Rankin, Gold, and Gooch (2006a) assessed the English language learning and vocabulary acquisition of language learners playing *Ever Quest 2*, a popular MMORPG. In their pilot-study, five ESL students, ranging from beginners to high-level learners, participated. They were asked to play the game for at least 4 hours per week for 4 weeks. They elicited information from the participants by post-questionnaire, interactions and chat logs, and administered a vocabulary test based on words used by the participants in their chat logs. Their results revealed that the digital game was not only beneficial for improving communication skills but also supported vocabulary acquisition. They found that vocabulary acquisition had taken place even for low exposure vocabulary items. Their participants accurately defined 35% of vocabulary items that they had been exposed to only once during their game playing sessions. The rate of acquisition of vocabulary item was greater the greater the frequency of interactions between players and non-playable characters (NPCs).

Yip and Kwan (2006) used online vocabulary games incorporated into two websites designed for language learning purposes by the University of Hong Kong, namely a professional word web and a university word web, to teach new vocabulary items. Participants were 100 freshman engineering students randomly divided into 6 groups. Groups A, B and C were experimental groups

and D, E and F, control groups. Each experimental group contained 15 or 16 participants and each control group 18 participants. The experimental group participants engaged in learning the new vocabulary item via web-based online games. The control group participants were taught the same vocabulary items through activity-based lessons. To familiarize the experimental group participants with the websites and teach them how to interact with them, they were asked to complete a series of online vocabulary tasks for the first 3 weeks. The controls, in turn, were given a list of words, derived from the websites, for rapid perusal. They were then asked to complete a series of vocabulary learning activities, such as drawing mind maps, matching definitions, drawing pictures, and designing a diagram, to help them develop vocabulary learning strategies. Using a pre- and post-test design, participants were administered 30 fill-in-the-blank questions in the first session and in the last session nine weeks later. Comparison of the pre- and post-test scores confirmed the superiority of vocabulary acquisition via digital games. Analysis of the post-test scores of both the experimental and control groups also showed that the experimental groups outperformed controls.

Hung (2011) analyzed an education-designed MMORPG, *ED-Wonderland*, to find out whether or not the educational game would achieve its goal of teaching English vocabulary items. He tested the vocabulary acquisition of participants in two different learning environments: in a multimedia instructional system and the digital game. Applying two design research cycles, he compared the two environments. He recruited 20 and 239 Taiwanese sixth and fifth graders, respectively, for his first and second research cycles. He compared the participants' motivation levels, learning performance and vocabulary retention between the two environments. After comparing the pre- and post-test results and mean scores of participants between the two environments, he found that the digital game better assisted vocabulary acquisition kept the participants more motivated and was more successful than the multimedia instruction system in teaching vocabulary in the long run, and that it achieved the goal it was designed for.

Muhanna (2012), in a pre- and post-test design, investigated the effect of digital games on vocabulary learning among 160 male and female Jordanian English learners in the 10th grade. The experimental group learned the new vocabulary items via a website game while the control group did not have the same opportunity. She reported that both the male and female experimental groups outperformed controls. She observed no significant gender difference the acquisition of vocabulary items via digital games.

Chen and Yang (2013) employed *Bone*, a commercial adventure digital game. to evaluate the effect of non-educational adventure digital games on learning English as a foreign language. In a pre-test-treatment-post-test design, 22 non-English speaking Taiwanese university freshman students were evaluated on their incidental acquisition of 20 target items via the digital game. After a pre-test, participants were given a game guide and a sheet of blank paper and asked to take notes. They played the game for one and half hours before

sitting a post-test. After comparing the final vocabulary post-test to pre-test results, they reported that the new vocabulary items were acquired incidentally during the digital gameplay They then conducted another study, in which they asked 35 college students for their perceptions of the digital game *Bone*. They gave their participants instruction on how to play the game. The participants were also asked to play both *Bone 1* and *Bone 2*. They had 16 weeks to finish the games. Through short written reports and a questionnaire, they collected data on the participants' perception. After analyzing the data, Chen and Yang (2013) reported that the students found the game fun and helpful in improving their language skills and motivation.

In a case study, Bakar and Nosratirad (2013) tested the effect of the game *SIMS 3* on adult vocabulary acquisition. They used interviews, observations, self-reports, and pre- and post-tests to find out how digital games aided 3 adult gamers (aged 22 to 30) in learning new vocabulary items. Analyzing the pre- and post – test scores and the qualitative data, they found that the digital games were facilitated vocabulary acquisition. They reported that, by playing the digital game, participants developed positive attitudes towards learning English. Moreover, the role of motivation was found to be very effective in boosting participants' vocabulary knowledge. They concluded that digital games, such as *SIMS 3*, could be advantageous instruments of vocabulary acquisition for adult learners who want to learn languages independently owing to such factors as freedom and control over performance.

Vahdat and Rasti Behbahani (2013) evaluated the effect of an adventure digital game on Iranian EFL learners' vocabulary acquisition. Participants were 40 intermediate learners assigned to four groups: a male experimental and control group and a female experimental and control group. The experimental groups encountered new vocabulary items via the digital game *Runaway: A Road Adventure* while the control group encountered the same vocabulary items via traditional vocabulary acquisition exercises. Comparing the final achievement tests by t-test, they reported that the experimental groups outperformed the control groups in vocabulary acquisition. They also mentioned a high positive correlation between gender and acquiring vocabulary via digital games, the male learners benefitting more form the digital game approach than females.

Janebi Enayat and Haghighatpasand (2019), in a pre-test/treatment/post-test design recruited 30 participants randomly assigned to either an experimental or control group. The experimental group played the adventure digital game *The Secret of Monkey Island- Special Edition* to practice target words that they had received in a word list before starting to play. The participants sat for an immediate receptive recall post-test and a delayed productive recall post-test. Between-group comparisons of test performances revealed that the experimental group outperformed controls in both post-tests.

Chen and Hsu (2019) recruited 66 college EFL students to play the digital game *Slave Trade*. In a quasi-experimental design, participants sat for a pre-test that evaluated them on their knowledge of receptive target vocabulary items and historical content. After playing the game, the participants were re-tested on the

same dimension of word knowledge. Comparison of the test results revealed that the digital game helped participants to learn both receptive words and content. They also found that receptive knowledge of words was acquired after at least 6 exposures while playing the digital game.

Finally, Sundqvist (2019) compared Swedish digital game players vocabulary knowledge level to non-gamers in a longitudinal study. She compared 1069 teenage (15-16 years old) Swedish speaking 9th graders to 16 Swedish (15-16 years old) 9th graders through 3 years of study. She asked the 16 participants to be actively playing digital games, especially commercial, among their out-of-school activities. She monitored their vocabulary development through questionnaires, English grades, productive and receptive vocabulary tests, and interviews and essays. Analyzing the results, she found that both time and type of gameplay were predicative variables that could predict vocabulary learning and development. However, the effect of type of gameplay disappeared when the data were analyzed in long-term analyses. She also found that gamers had developed vocabulary better than non-gamers. She finally concluded that frequency of gameplay, time spend for playing, and the commercial digital games could be considered important for vocabulary learning and development.

Digital gameplay has also been found to boost the acquisition of specific/technical as well as no-technical words. For example, Fotouhi-Ghazvini, Earnshaw, Robison, and Excell (2009), found that digital games assisted the acquisition of field-specific vocabulary items. They designed a mobile phone game package to aid incidental technical vocabulary acquisition for the college students majoring in computer science. Participants were 15 students assigned to 3 groups of 5 participants each. The first group read a text containing the target words. The second group used a dictionary to memorize the target words. The third group played *MOBO City*, a digital game designed by the researchers. Analysis of the results revealed the incidental acquisition of technical computer science vocabulary items like PCI, AGP, CPU etc. Jasso (2012), in turn, investigated the consequences of using a non-academic digital game on vocabulary acquisition in the EFL classroom. Participants were 14 beginning level EFL learners in a French university. She tested participants' acquisition of clothes-related vocabulary items by playing *SIMs,* a commercial digital game called. Her productive post-test results showed that the experimental group participants, who had encountered the vocabulary items via the digital game, outperformed the control group participants, who had encountered the same group of words via magazine pictures. She attributed the experimental group's better vocabulary acquisition via the digital game to higher motivation.

Having considered the various digital game genres used in the studies reviewed above, I came to the conclusion that the adventure genre would be a suitable choice for my study. This, for various reasons: first, because of its popularity. adventure digital games are preferred by 62.1% of gamers (Zin, Jaafar, & Yue, 2009). Second, adventure games can be both entertaining and educational simultaneously. Third, from the standpoint of narratology and ludology, both story and gameplay are combined in adventure games. In other

words, they are not separate realities in adventure digital games but also determining factors during play (Ang, & Zaphiris, 2006). Moreover,

> in adventure games there are very complex environments i.e. microworlds, with no deterministic problem representation. Adventure games use intrinsic motivation. Intrinsically motivating games incorporate learning activity in a virtual world. Game characters have to solve a certain problem and can proceed further only after solving the problem. In this case the problem is part of the game and players are motivated to provide a solution in order to continue with the game. (Pivec, Dziabenko, and Schinnerl, 2003, p. 218)

Finally, compared to other genres, adventure digital games can outperform other genres in gaining learners' attention, informing learners about the educational objective, simulating recall of prior learning, presenting stimuli, providing learning guidance, eliciting performance, providing feedback, assessing performance, and enhancing retention, all of which, according to Gagné, are elements of a successful educational setting (Becker, 2008). Thus, a commercial adventure digital game was seen a suitable selection for this study.

How digital games enhance vocabulary acquisition has also been discussed. That is, the cognitive-motivational effect of internal game elements has mostly received attention in seeking to explain this effect. In general, digital games present content via rich images, animations, videos, visuals and audios, thereby providing opportunities for high frequency of occurrence, variation in mode of presentation and authentic contexts, which are known to be effective factors in vocabulary acquisition (Sundqvist & Sylvén, 2012; Hwang & Wang, 2016; Zhonggen, 2018; Janebi Enayat & Haghighatpasand, 2019).

To test this idea, Ebrahimzadeh and Alavi (2016), in a co-relational study, evaluated the effect of various internal factors on the escalation of enjoyment, leading to a state of flow and, finally, motivation. They randomly assigned 136 participants (12-18 years old) to two groups, namely, players and watchers. To expose themselves to the target vocabulary items, participants were asked to either play or watch a digital game. Data were obtained using an e-learning enjoyment scale, field notes and vocabulary post-tests. Although no difference was observed between watchers and players in vocabulary acquisition, the internal factors that support enjoyment, flow and motivation showed a positive correlation with vocabulary acquisition. They concluded that, in any digital game, the presence of internal motivational factors, especially, challenge, immersion, autonomy and knowledge improvement, could predict its efficacy in enhancing vocabulary acquisition because the digital game would increase enjoyment, lead to a state of flow and keep the gamers motivated.

The motivational effect of the internal elements of digital games on vocabulary acquisition was also supported in a meta-analysis by Chen, Tseng, and Hsiao (2018). They found that game design elements could play a key role in inducing flow in DGBLL and vocabulary acquisition. Moreover, they reported that adventure digital games would be more effective for vocabulary acquisition than non-adventure digital games, since they could provide an optimal level of an important internal element, namely challenge.

However, as mentioned before, the internal elements of digital games could have a bipolar effect on vocabulary acquisition, that is, motivational and cognitive. Although motivational part of effect was found effective on vocabulary acquisition, the cognitive part of the effect was found conditional and, sometimes, hindering. The effect of interactivity, as one of the internal elements of digital games that causes cognitive effects, on vocabulary recall was found conditional (deHaan, Reed, & Kuwada, 2010) even though, Zhonggen (2018) found that interactivity-prone games would facilitate vocabulary acquisition more efficient than less interactivity-prone digital games. deHaan, Reed, and Kuwada (2010) investigated the effect of interactivity with a musical digital game on second language vocabulary recall. They classified 80 randomly-selected Japanese university undergraduates into two groups of watchers, who were only watching a digital game that were played, and players, who were playing the digital games. The players were asked to play a musical, rhythm game. And the watchers were asked to watch the gameplay. A t-test comparison of two-week delayed vocabulary recall post-tests revealed that watchers, with low level of interactivity, achieved more vocabulary items than game players. They concluded that high level of interactivity with digital games leads to cognitive overload and, consequently, fewer vocabulary item retention and recall. This finding matters for vocabulary acquisition because vocabulary acquisition is a complex and multidimensional process that can be affected by multiple factors.

One of the factors that can affect vocabulary acquisition is the cognitive nature of the task (Hulstijn & Laufer, 2001). In a study on the suitability of mobile digital games for incidental vocabulary acquisition, Reynolds (2017) took an internal element perspective on DGBL to find cognitive-motivational factors that contribute to digital game-based vocabulary learning tasks. He asked 92 Taiwanese undergraduates and graduates to play the social game *Draw Something*. The results of his post-performance questionnaire revealed that digital game task-induced involvement load, as a prevalent cognitive-motivational factor, played a significant role in the gamers acquisition of new vocabulary items. He found, in an ANOVA analysis of participants' perceptions of interactions in digital games, that the gamers involved themselves in components of involvement load, such as search, then need, and finally evaluation. He concluded that the involvement load hypothesis should be considered when choosing a digital game for language learning. Furthermore, he states that "how useful one particular game can be for inducing vocabulary acquisition depends on the amount of task-induced involvement" (Reynolds, 2017, p. 482). Finally, he recommended the involvement load hypothesis as a precise touchstone for analyzing digital games and their suitability for DGBLL purposes. Although he emphasized the importance of digital game-based task-induced involvement load, he did not evaluate its effect on vocabulary acquisition, leaving a gap in the digital game-based vocabulary acquisition literature.

3.6 Research Questions:

In regards with the importance of internal elements perspective, findings of deHaan, Reed, and Kuwada (2010), emphasis of Reynolds (2017), and importance of the effect of task features on vocabulary acquisition (Hulstijn & Laufer, 2001), evaluating the effect of digital game-based task-induced levels of involvement load seems necessary and promising. Therefore, as there are not many studies conducted to point and to fulfill this gap in literature, this study is an attempt to evaluate the effect of digital game-based vocabulary learning task, inducing different levels of involvement load, on the acquisition of new vocabulary items both qualitatively and quantitatively. Moreover, due to multidimensionality of vocabulary acquisition (Ringbom, 1987; Nation, 2001), this study evaluates the effect of digital game tasks on the acquisition of dimensions and scopes of word knowledge, such as receptive (recognition and recall), and productive (recognition and recall) knowledge. Therefore, in order to achieve the aim of this study, I will answer the following questions:

1. What is the effect of the digital game, in different levels of involvement load, on the acquisition of target vocabulary items?
    1.1. Does playing the digital game make a significant difference in the participants' performance, in their *productive* tests, before and after playing the digital game?
    1.2. Does playing the digital game make a significant difference in the participants' performance, in their *receptive* tests, before and after playing the digital game?
2. Which dimension and scope of word knowledge, either receptive (recall/recognition) or productive (recall/recognition), are acquired significantly better after completing digital game tasks in different levels of involvement load?
3. Does interacting with the digital game tasks, in different levels of involvement load, make significant differences in vocabulary acquisition?

The first research question is a general question that I feel should be asked in every language learning-related digital game study. The answers to this question should clarify whether the treatment, i.e. the digital game task, has been effective enough to generate the formation of new constructs, in this instance the acquisition by gamers of target words. Moreover, by dividing the first question into two parts to facilitate more detailed analysis and, in the second question, scrutinizing the dimensions and scopes of vocabulary knowledge, I would be able weigh the effect of the digital game tasks on each of the two central aspects of vocabulary acquisition, i.e. productive (recognition/recall) and receptive (recognition/recall) knowledge. The last question, which addresses the main aim of this study, was designed to identify issues that are missing in the digital game-based vocabulary acquisition literature. Finally, to deepen knowledge on the phenomenon, I applied both qualitative and quantitative approaches to the data.

4 METHODOLOGY

The main purpose of this study was to evaluate the effect of digital game tasks inducing different levels of involvement load on the acquisition of target vocabulary items. In other words, I not only investigated the possible effects that each of three levels of involvement load may have on the acquisition of target vocabulary items but also at what level of involvement load new vocabulary items would be most effectively acquired in digital game-based vocabulary acquisition tasks. However, this goal is not easy to achieve by looking at test scores and quantitative data alone. Thus, as mentioned earlier, I also obtained qualitative data by applying a think-aloud protocol and conducting an exit interview.

In this chapter, I report how the participants were selected, describe the instruments, materials and study design used for collecting the data, and detail the research process. In the data analysis section, after reviewing content analysis in detail, I explain how the concurrent think-aloud data were analyzed. I also discuss and justify my choice of content analysis as the preferred method of data analysis for this study.

4.1 Participants

Concerning the rating of a digital game used in this study and simplicity of the task, participants had to be no younger than 13 years old and had to be able to comprehend simple texts of their tasks. Moreover, it was important to recruit participants who were sufficiently talkative and eager to communicate with their partners when performing language tasks. After finding a cooperative private language institute in Iran in summer 2018, and consulting with the head of the institute, I found that the pre-intermediate language learners in the institute were mostly aged 12 years or older. There were 5 classes of pre-intermediate language learners whose English proficiency had been assessed by the institute's own

internal language proficiency test[3]. I visited all 5 classes and described the study and its purposes, including what participants would do, and asked for volunteers, as participant interest is a notable factor in determining the quality of collected verbalized data in any think-aloud study (Charters, 2003). Fortunately, many language learners were interested in participating. Eventually, 30 Iranian male (n=14) and female (n=16) language learners ranging in age from 13 to 15 years were randomly selected to participate in the study. All participants were native speakers of Persian studying English as a foreign language in the participating private language institute in Iran in summer 2018. They had started learning English at age 11 in their junior high schools before registering in the language institute. They attended classes in the institute in the evenings and after school three times a week. At the time of recruitment, they had been studying English in the institute for one year. Apart from their English homework, they reported that they rarely studied English in their free time, although they might have encountered English in public media such as TV, movies, digital games, internet etc. All reported that, while they had played digital games them in their free time for fun, they had never used them for the purpose of language learning. Although their proficiency level had been evaluated by the institute as intermediate, I tested their vocabulary knowledge, using Nation's Vocabulary Size Test (Appendix A), to ensure that they were eligible for participation in my study. Their vocabulary size test scores ranged between 2000 and 3300 word families.

To comply with research ethics, formal consent to participate in the study was requested from both the participants and their parents. In giving their consent, they mentioned that their first name could be reported provided that their last names remained anonymous. Participants were next asked about their gameplay habits, as persons with more experience in playing digital games have more expertise in acquiring new skills in novel digital game contexts (Schrader & McCreery, 2008). All participants reported habitually playing digital games in their free time. They reported playing different types of digital games on consoles, PCs, tablets, and mobile phones for at least 1 hour a day. The boys mostly played action and adventure games such as *Grand Theft Auto, Call of Duty,* and *Counter Strike*. The girls mostly preferred playing adventure, puzzle, and casual games on either their tablets or mobile phones. None of the participants had ever previously played *Haunted Hotel: Death Sentence Collector Edition*, the game selected for this study. However, when I asked them about their digital game playing habits it emerged that they had experienced games similar to *Haunted Hotel* at least once.

To collect the concurrent think-aloud data, I selected a sub-sample of twelve participants after consulting with the teachers of the pre-intermediate groups. Specifically, I randomly selected two pairs from each of the three experimental groups, namely A, B, and C. All were over age 13 and were considered to be both

[3] I was informed that the maximum score of the internal proficiency test was 20. Scores between 0 - 5 denote beginner, 6 - 10 pre-intermediate, 11 - 15 intermediate, and 16 - 20 advanced level.

active and communicative when performing language tasks during their language courses. Their native language was Persian, and they used English as a foreign language in their English classrooms. Except for English, they had studied most of their school subjects in Persian. Their demographical information is presented in table 8.

TABLE 8 Demographics of participants selected for the think-aloud data collection

Name	Gender	Language	Age	Vocabulary Size
Shirin	Female	Persian	15	3300
Mähdieh	Female	Persian	15	2500
Moein	Male	Persian	14	2400
Shadi	Male	Persian	13	2200
Minoo	Female	Persian	15	2100
Äli	Female	Persian	14	2100
Soheil	Male	Persian	13	2200
Artin	Male	Persian	13	2400
Tara	Female	Persian	15	2300
Negar	Female	Persian	14	3100
Äli 2	Male	Persian	14	2800
Älireza	Male	Persian	13	3000

4.2 Instruments

To measure participants' linguistic knowledge and to explore their cognition, I used three instruments: the vocabulary size test (Nation & Beglar, 2007), achievement tests (both receptive and productive), and an interview. These instruments are described in detail below.

4.2.1 Vocabulary Size Test

To ensure the homogeneity of the participants in their vocabulary knowledge, I measured their vocabulary size, or breadth of word knowledge (Nation, 2001), by a standard test known as the vocabulary size test (Nation & Begler, 2007). This test is widely accepted by both applied linguists and language teachers (Schmitt, 2010; Nation & Webb, 2010; Read, 2000; Schmitt & McCarthy, 1997). The vocabulary size test was a suitable choice here, as it is a reliable test for "selecting individuals displaying specific levels of vocabulary knowledge for particular educational experiences" (Beglar, 2010, p. 102) as well as for either grouping or choosing suitable participants in vocabulary acquisition empirical studies (Elgort, 2012). The test can also determine if the selected samples are suitable for the study (Schmitt & McCarthy, 1997).

The vocabulary size test is a multiple-choice test with 140 stems or questions. Each stem is a randomly selected vocabulary item, or word, followed by an example that shows the use of the word in a real-life simple sentence. Four

alternative responses, which are suggested definitions in English, are offered for each stem. Only one of the alternatives closely matches each stem and is deemed the correct answer. The selected stems are divided into 14 groups by frequency of occurrence; the higher the group ranking, the lower the frequency of the words in that group. In other words, test takers may find more familiar words in the first 10 stems, i.e. group 1, rather than the final 10 stems, i.e. group 14. Scores for the vocabulary size test are calculated by multiplying the sum of correct responses by 100. The result shows the test-taker's proximate receptive knowledge of word families. In other words, the test taker knows N word families, i.e. no fewer than his score in the vocabulary size test.

TABLE 9 An example of a multiple-choice test item in the vocabulary size test

POOR: we are **poor**
a. Have no money
b. Feel happy
c. Are very interested
d. Do not like to work hard

I adapted a bilingual vocabulary size test for this study. In so doing, I translated only the alternatives into Persian (Nation, 2012). When translating, I was especially careful regarding cognates (Beglar, 2012). Because the Persian language is an Indo-European language and shares many similarities, especially in vocabulary, with other members of this family, such as English (Yule, 2014), I either used another Persian synonym, if it was available, or explained the definitions in short sentences. For example, in the case of the word *nun*, which is also used in Persian but with a slight difference in pronunciation, I explained it in a short sentence. In the case of the word *microphone*, which is considered a borrowed word and is used as it is, I offered the synonym "بلندگو".

The test was translated and administered in bilingual format for two important reasons. First, as Beglar (2012) argues, "in a foreign language context, participants who know the conceptual meaning of an L2 word being tested may be disadvantaged if they misunderstand a definition containing a noun/verb phrase […], owing to their insufficient knowledge of grammar or syntax" (p. 254). Moreover, she states that bilingual tests developed for measuring the vocabulary knowledge of language learners are considered less stressful because the context of these tests is supposedly more familiar to the test takers. Moreover, in a familiar context, test takers may also be willing to attempt less familiar test items. Thus, we can infer that precision in measuring and comfort in performing might be valuable properties of bilingual tests. Therefore, the bilingual vocabulary size test probably served the purpose of this study reasonably well and produced a vocabulary size index that approximated the participants' true receptive vocabulary knowledge. The Vocabulary Size Test is given in Appendix A.

4.2.2 Achievement Tests

For study the effect of the independent factors, i.e. task-induced involvement load and the digital game, on the acquisition of target vocabulary items, it was necessary to evaluate the participants' task performance. Webb (2005 as cited in Nation & Webb, 2010) emphasizes that deeper analysis of vocabulary knowledge, in any vocabulary learning study, is necessary as it provides the researcher with more clues in deciding whether or not the target vocabulary items have been acquired. Hence, to obtain a more profound insight of participants' knowledge of the target words, I designed two types of achievement test: a receptive test and a productive test. The tests, described in section 2.3, were bilingual for the reasons discussed by Beglar (2012). Moreover, the bilinguality of the tests, aside from its benefits, made it easier for me to measure the two dimensions and scopes (receptive/productive – recognition/recall) of word knowledge. However, it is important to note that the use of bilinguality in vocabulary testing could have a learning effect, a factor that must be considered comparing participants' performance in sequentially administered tests (Mondria & Wiersma, 2004).

The receptive vocabulary knowledge test measured participants' ability to recognize and recall the Persian definitions of the target vocabulary items encountered during their engagement with the digital game (receptive recognition/recall) (Nation, 2002; Read, 2000). The receptive test comprised three sections. In the first section, the test takers' recognition of the definitions of the target words was measured. They were asked to select the English words (form) from the list shown on the left that match the Persian definitions (meaning) shown in random order on the right. The total number of English words was 12 and the Persian definitions were 6 (Table 10).

TABLE 10 Receptive-recognition example question in first section

1. Debris	
2. Shack	_____ چِفت در
3. Latch	_____ خُرد و ریز
4. Hook	_____ قلاب
5. Skull	
6. Matches	

The second section comprised 4 multiple choice items. Each item featured one of the target words (bolded), followed by 4 alternative definitions in Persian. The test takers were asked to select the definition that best corresponded to the target word. Both the first and second section of the test measured receptive - recognition knowledge of the target words (10 questions).

TABLE 11 Receptive-recognition example question in second section

Portrait means:			
a) قاب عکس	b) فَرش	c) کبریت	d) جمجمه

The final section of the test contained the remaining 10 target words. This section also evaluated participants' receptive - recall knowledge. The test takers had to translate the words, which were not used in sentences, into Persian by recalling their definitions.

The productive vocabulary knowledge test evaluated the extent to which the participants had acquired productive knowledge of the target vocabulary items. Specifically, the test was administered to determine to what extent participants could recognize and recall the target words, which could be either spoken or written (Nation, 2002). The productive vocabulary knowledge test was divided into 3 sections. The first section tested the participants' knowledge of the form-meaning link. The task was to select the Persian definitions from the list shown on the left that match the English forms shown on the right (Recognition). There were 18 Persian definitions and 10 English words. An example in shown in table 12.

TABLE 12 Productive-recognition test example in first section

1. قلاب	
2. فَرش	_____ Fragments
3. چفت در	_____ Latch
4. گنجه	_____ Debris
5. خُرد و ریز	
6. تیکه خُرده	

The second section contained 4 fill-in-the-blank items. The test takers had to read a sentence and fill in the blank with an appropriate word (Recall). To retrieve the correct word forms, they had to consider both the context and clues, which I intentionally included. For example, in the sentence "*I got a big fish with a long* ***ho***_____.", the test takers had to understand the context, which is about fishing; furthermore, they had to retrieve the full form of the word that started with "*ho"* and fill in the blank with the rest of the word *hook*. Finally, in the third section, they had to recall 6 English forms through translation (Recall). The Persian definitions were given on the left and the test takers had to translate them into their English equivalents. no clues, such as initial/middle/final letters, or sentences were given in the last section. Both the receptive and productive tests had a maximum score of 20. The tests can be found in Appendix B.

4.2.3 Interview

Qualitative as well as quantitative data were collected in this study. To enrich the qualitative data, collected through the think-aloud protocol, and also to elicit retrospective data, I interviewed selected think-aloud participants (N = 12) immediately after they finished their tasks. This type of interview, known as an "exit interview" (Charters, 2003), must not be confused with retrospective think aloud, as discussed earlier (Section 2.7.1.2). The interview was retrospective in that I administered it immediately after the participants had completed the

concurrent think-aloud task to elicit more information to support the collected concurrent think-aloud data.

The interview was conducted in Persian and only concerned the think-aloud participants. In the first section of the interview, I asked the participants about their background and whether they had any experiences of learning vocabulary via digital games. In the second section of the interview, they were asked about their learning experience through the digital game used in this study. In this section of the interview, I was seeking not only to evoke participants' thought processes when engaging with the target words but also to find out what strategies they used to overcome challenges they encountered either in playing the digital game or in the text where the target words were embedded and, finally, what factors in both the digital game and the text that either hindered or aided their progress. The interviewees were given the opportunity to express their own ideas as well. Moreover, they were asked what factors they would have included or excluded in a digital game-based vocabulary acquisition context if they themselves were an educational designer. Finally, it should be mentioned that the cognitive load-related questions and ideas used in my interview were adapted from the *Cognitive Load Subjective Experience Questionnaire* (deHaan, 2008). The list of interview questions is given in Appendix H.

4.3 Materials used in this study

I used two materials in conducting this study. these were the digital game and a game guide. Engaging with these two materials, participants were expected to learn 20 target vocabulary items. The combination of both the digital game and the game guide led me to administer three different types of task one inducing a low, one a moderate and one a high level of involvement load. Below, I describe the digital game, the booklets, and how the tasks, inducing three different levels of involvement load, were created.

4.3.1 The Digital Game

The digital game that I chose for this study was *Haunted Hotel: Death Sentence Collector's Edition*. The game was rated T, meaning suitable for teenagers (age 13 and above) by the Entertainment Software Rating Board (ESRB). The ESRB describes the game as follows:

> This is a […] game in which players help investigate a series of crimes. Players solve puzzles and search for specific items among screens of jumbled objects to solve each case. [The] game contains dark thematic elements (e.g., references to murder, kidnapping) and brief instances of violence: a man shot at by a sniper. ("Haunted Hotel," 2017a)

The game was developed and published by the Big Fish Game Studio in 2015. The studio is world-widely known for its adventure, hidden-object, and puzzle-solving digital games. The digital game genre is both adventure and hidden object. Gamers control the game by mouse clicks only.

FIGURE 8 Haunted Hotel: Death Sentence, the Game Cover

In the game, the gamer occupies the role of a private detective who is trying to find his friend. After finding a mysterious letter from his friend, he travels to a forsaken hotel where he encounters a series of events. The game publisher describes the story follows.

> Together you and your friend James have solved all kinds of supernatural mysteries together, but this case may be his last. Late one evening, you receive a letter written in James's own handwriting, claiming that he has died and that the Holy Mountain Hotel holds the answers. Racing off to the hotel, you discover it's completely abandoned. But is it really? You'll have to dig deep to uncover the murderer responsible for taking justice into his own hands! Can you save James and escape with your lives, or will The Holy Mountain Hotel become your early grave? ("Haunted Hotel" 2017b)

The game contains a series of events that must be solved by finding specific objects. In psychological terms, the events are problem-solving tasks that are inter-related (Pivec, Dziabenko, & Schinnerl, 2003; Becker, 2008). If gamers cannot solve the first problem, the first event does not happen and, consequently, the second problem does not appear, that is to say, the game does not progress. Gamers can find the required objects in one of the following ways: by searching the game areas, by combining collected objects, or by solving issues in previous events. Likewise, the objects must be either used in a specific area or place in the

game or combined with other objects to create another useful object. For instance, the following figures show two examples of how objects can be used in the game.

FIGURE 9 The gamer is trying to break into the shack by a stone from his inventory (Object on the place).

FIGURE 10 A gamer has found the glue and broken door-handle fractions. He applies glue to attach them into a new object, i.e., door handle (object combination).

Although serious games have been designed for educational purposes, I used a commercial digital game because it "enables students to use a high-end and attractive product", which better attracts teens' interest (Reinders & Wattana, 2012). Moreover, commercial digital games are the main target of digital game consumers rather than educational digital games (Zin, Jaafar, & Yue, 2009; Chen & Yang, 2013). In addition, it was easier to find a commercial than serious educational digital game that satisfied the background definition of a digital game and also fully served the purpose of this study. Moreover, since commercial

digital games outnumber serious games, I had more options from which to choose.

4.3.2 The Game Guide and Task Design

The second material used in this study was a game guide. The game guide was downloaded from the website of the official developer, Big Fish Games. To integrate the game guide into this study, it was revised in various ways. First, most of the picture guides, which were related to the story and main events of the game, were replaced with short sentences. Two pictures were left in because they were helpful for solving two time-consuming mini-games, which were integrated for entertainment purposes and unrelated to the main story events; in addition, retaining those two pictures could have reduced the time taken to do the task and thus have stopped the participants from losing track of the story, the game, and the task. The game text was then simplified. The demands on comprehension of the context, and the level of difficulty the text were reduced by removing less frequent words, restructuring sentences to create imperative forms, breaking them down into smaller imperative sentences and numbering them. Finally, the target words, which learners were supposed to learn, were presented in bold font as this is considered an appropriate strategy for attracting learners' attention and concentrating their focus on the words (Schmitt, 1997).

A total of 20 target words was selected from the game guide. These were all nouns, presented in bold font, and relatively infrequent words. They were the names of objects like *magnifier (K4)* and *skull (K5)* mentioned in the game. When selecting the target words, three criteria were followed: first, they had to be nouns; second , the gamer would not able to progress further in the game without knowing those words; and third, the words had not been repeated more than twice in the game guide. I selected nouns as previous studies have reported that nouns are easier to acquire than other word classes, and their priority in the order of acquisition has been empirically confirmed (Kiss, 1973; Ellis & Beaton, 1993b; Ellis, 1995, 1997; Nation, 2002). Besides, limiting learning to nouns would also reduce the burden of word difficulty in the learning process (Ellis, 1995; Söckman, 1997). Furthermore, in one of my recent studies, I found that the acquisition of nouns was quicker than that of other lexical word classes in a digital game-based vocabulary learning context (Rasti Behbahani, 2017).

In this study, repetition was a control variable, as the key role of repetition in learning new words is very well recognized and its importance emphasized by vocabulary learning researchers (Nation, 2002; Webb, 2007; Schmitt, 2008, 2010a, 2010b; Nation & Webb, 2010). Thus, to study the effect of task-induced involvement load and the digital game alone on the acquisition of the selected target words, they were not repeated more than twice.

The level of involvement load was also adjusted by modifications made to the game guide. To adjust the levels of task-induced involvement load, two types of glosses were integrated and a technique, called pre-teaching new words, was applied. Integrating single glosses, meaning-inferred glosses, and using pre-

teaching target words technique enabled me to modify the game guide for the purpose of inducing three different levels of involvement load.

Task-induced involvement load was measured by the Technique Feature Analysis (TFA) checklist owing to its precision in indexing the level of involvement load that each vocabulary learning task may induce (Nation & Webb, 2010).

Three tasks with low, moderate, and high levels of involvement load named A, B, and C tasks, respectively, were set. In task A, the game guide was modified to provide single-word glosses, the 20 target words were presented in bold, and the Persian definition of each word was placed in front of that on glosses. Applying the TFA checklist, task-induced involvement load index for game guide/task A was 7.

TABLE 13 Checklist for Technique Feature Analysis of the Group A Task

Criteria	Scores
Motivation	
Is there a clear vocabulary learning goal?	1
Does the activity motivate learning?	1
Do the learners select the word?	0
Noticing	
Does the activity focus attention on the target words?	1
Does the activity rise awareness of new vocabulary learning?	1
Does the activity involve negotiation?	0
Retrieval	
Does the activity involve retrieval of the word?	0
Is it productive retrieval?	0
Is it recall?	0
Are there multiple retrievals of each word?	0
Is there spacing between retrievals?	0
Generation	
Does the activity involve generative use?	1
Is it productive?	0
Is there a marked change in context that involves the use of other words?	0
Retention	
Does the activity ensure successful linking of form and meaning?	1
Does the activity involve instantiation?	0
Does the activity involve imaging?	0
Does the activity avoid interference?	1
Maximum score	**7**

For the task B, the game guide was modified by the addition of multiple-choice meaning-inferred glosses and the 20 target words were bolded. However, in this task, three different Persian definitions were presented for every target word in front of that on marginal glosses of the game guide. Linguistically speaking, glosses were given in task A and multiple-choice glosses were given in Task B. I also limited the number of repetitions of the Persian definitions. Accordingly, the

task B participants were exposed to each definition three times. The index of task-induced involvement load in task B was 9.

TABLE 14 Checklist for Technique Feature Analysis of the Group B Task

Criteria	Scores
Motivation	
Is there a clear vocabulary learning goal?	1
Does the activity motivate learning?	1
Do the learners select the word?	0
Noticing	
Does the activity focus attention on the target words?	1
Does the activity rise awareness of new vocabulary learning?	1
Does the activity involve negotiation?	1
Retrieval	
Does the activity involve retrieval of the word?	1
Is it productive retrieval?	0
Is it recall?	0
Are there multiple retrievals of each word?	1
Is there spacing between retrievals?	0
Generation	
Does the activity involve generative use?	1
Is it productive?	0
Is there a marked change in context that involves the use of other words?	0
Retention	
Does the activity ensure successful linking of form and meaning?	0
Does the activity involve instantiation?	0
Does the activity involve imaging?	0
Does the activity avoid interference?	1
Maximum score	**9**

In the task C, glosses were not employed. The 20 target words were pre-taught by giving the participants a word list that contained the Persian definition for each target word and an example of each word in a simple sentence after each word. The example sentences were genuine and were not related to the game or the game guide. In the task C game guide, as in other two groups, the 20 target words were bolded. Task C had the highest task-induced involvement load index, i.e., 12.

TABLE 15 Checklist for Technique Feature Analysis of the Group C Task

Criteria	Scores
Motivation	
Is there a clear vocabulary learning goal?	1
Does the activity motivate learning?	1
Do the learners select the word?	0
Noticing	
Does the activity focus attention on the target words?	1
Does the activity rise awareness of new vocabulary learning?	1
Does the activity involve negotiation?	1
Retrieval	
Does the activity involve retrieval of the word?	1
Is it productive retrieval?	0
Is it recall?	1
Are there multiple retrievals of each word?	1
Is there spacing between retrievals?	1
Generation	
Does the activity involve generative use?	1
Is it productive?	0
Is there a marked change in context that involves the use of other words?	0
Retention	
Does the activity ensure successful linking of form and meaning?	1
Does the activity involve instantiation?	0
Does the activity involve imaging?	0
Does the activity avoid interference?	1
Maximum score	12

Participants were given the game guide to prevent the negative effects of extensive interactivity. A previous study found that high levels of interactivity, or extensive interactivity, with a digital game can result in negative cognitive load, which probably leads to malfunction in learning new concepts (Mayer & Moreno, 2003; Sweller, 2010), especially in recognizing, recalling, retrieving and retaining vocabulary items (deHaan, 2008). Therefore, in light of these previous findings, especially those of deHaan (2008), the extent of interactivity was controlled in the present study design by implementing a signaling technique, that is, "providing cues to the learner about how to select and organize the material" (Mayer & Moreno, 2003), in this instance the game guide.

4.4 Procedure

In conducting this study, participant selection was the first step. After obtaining an official letter, requesting cooperation and clarifying the aim of the data collection, from the University of Jyväskylä, several schools, educational organizations, and foreign language institutes in Iran, both public and private, were contacted. The search ended with a private foreign language institute that

showed interest in the study and promised cooperation. As described in section 4.1, a total of 30 participants were randomly selected for this study. All the participants were then requested to sit for a vocabulary size test to evaluate the breadth of their vocabulary knowledge and pre-tests, designed to test their knowledge of the selected 20 target words. Their vocabulary size scores ranged between 2 000 and 3 300, indicating that the participants' English vocabularies were fairly homogeneous in size. The pre-test results also revealed that the target words were all new for the participants. Finally, I randomly assigned the participants in pairs to the three groups A, B, and C. Thus, each group contained 10 participants, i.e., 5 pairs, who played the digital game while reading the game guide. Furthermore, 2 of the 5 pairs in each group were selected for the qualitative data collection.

In general, while all participants, regardless of group followed the same instructions, the tasks assigned to each group induced a different level of involvement load. Participants were expected to read their game guide texts and cooperate in their gameplay. Additionally, the think-aloud pairs had to verbalize the ideas and thoughts that passed through their minds while engaging in their tasks. Because the gameplay was cooperative, the participants were expected to discuss their ideas and thoughts with their partners while playing the game. They were asked to do the tasks in pairs because such "tasks are an integral part of second language learning" (Cohen, 1987, p. 90) and are ubiquitous in language learning classrooms. Think-aloud partners' voices, actions, and gameplay footage were recorded by audio recorders and cameras, and their gameplay was recorded in the computer using the screen capture software *Fraps*.

In applying the non-metalinguistic concurrent think-aloud method for qualitative data collection, the recommendations of previous studies were followed. Accordingly, the think-aloud pairs had a warmup session (Ericsson & Simon, 1987), a general and broad instruction or protocol, and a reminder (Leow & Morgan-Short, 2004). The think-aloud pairs were invited to the warmup session two days before their main task. The warmup session was a simulation of the main task. During the session, I introduced the digital game, how to play it, and how to complete the task. I loaded one of the later scenes in the game, unrelated to their task or this study, and offered them an opportunity to practice by playing. During this session, they were expected to learn how to control the game, cooperate with their partners, become accustomed to being recorded, verbalize their thoughts, and discuss their ideas with each other. The session lasted 30 minutes and they played on six Windows™ laptops. After the warmup session ended, I delivered the target word list to the group C participants and asked them to memorize all 20 words beforehand and return the list to me the following day. Two days after the warmup session, all participants performed the main task. To reduce possible protocol effects in the case of the think-aloud pairs, they were instructed to "verbalize whatever goes through your mind when you are playing, and it is desirable to discuss these thoughts with your partner". I used my presence as a reminder. I informed them in the following words: "I will be sitting behind the door and will come in randomly as a reminder or I will

give the signal: Keep Talking. If you are silent or have forgotten to talk or discuss, please remember that you have to tell your thoughts and discuss them with your partner". I left the room but remained nearby, as the presence of the researcher in the same space could impair the validity of the data (Ericsson & Simon, 1993). Moreover, "if the subjects are too much aware of the presence of the researcher, they may try to produce 'socially acceptable data' "(Cohen, 1987, p. 89). I visited the think-aloud participants almost every 7 to 10 minutes.

All the data collection sessions were held over three days in summer 2018. On the first day, group A performed their task, followed on the second day by group B, and on the third day by group C. The think-aloud pairs were seated in separate rooms because their voices needed to be recorded. The other participants were all in the same room. They spent between 70 and 90 minutes on their tasks. Immediately after the think-aloud participants had finished their task, I interviewed them individually for about 20 minutes each. The interviews were audio-recorded. The participants were not forewarned about the exit interview. Ericsson and Simon (1980) state that participants must not be told explicitly about an exit interview as this could affect their cognitive processes during the main task performance. Finally, 3 weeks after the main task, the delayed post-posts were administered in the same setting. Schmitt asserts that "a delayed posttest of three weeks should be indicative of learning which is stable and durable" (2010, p. 157). It should be mentioned that, in administering both the pre- and post-tests, the receptive test preceded the productive test (Nation, 2002).

The design of this study resembles a semi-experimental research design, i.e., a pre-test/treatment/post-test design. Moreover, the qualitative data were also collected through the concurrent think-aloud protocol. In other words, in this study, both qualitative and quantitative methods of data collection were implemented. There were two reasons for this: first, to elicit information about online cognitive processes that are evoked by different levels of involvement load induced by digital game-based vocabulary acquisition tasks, and second, to assist the researcher in finding out to what extent the target words had been acquired after the experimental treatment, that is, after playing the digital game. Thus, this study can be described as a mixed methods study designed to reap the advantages of both qualitative think-aloud data and quantitative statistical results.

4.5 Data Analysis

The data were both qualitative and quantitative. I applied three statistical methods in analyzing the quantitative data and used qualitative content analysis in analyzing the qualitative data. These methods are described in detail below.

4.5.1 Quantitative Data Analysis

To analyze the quantitative data, which were measures of the effect of the digital game on the acquisition of the target vocabulary items at three different levels of involvement load, statistical formulae were implemented using the *Statistical Package for Social Sciences* (SPSS), version 23.

The methods for analyzing the quantitative data were, for obvious reasons, recommended to be selected from among the non-parametric methods of data analysis. First, the study population, comprising 30 participants was small and each of the three study groups contained only 10 participants. Second, a small study population means a small data distribution, which in turn would render the outcome of parametric methods of analysis less reliable than that of non-parametric methods. Finally, non-parametric methods of analysis allow the removal of unintended outlier effects, that is, the effects of the means that are significantly different from the main distribution of the data (Brown, 1988; Hall, 2015). In other words, unlike parametric methods, data analysis using non-parametric methods of analysis does not rely on distribution patterns. Thus, to ensure the reliability of outcomes and to remove possible unexpected problems induced by the distribution of a small population, non-parametric methods of data analysis were selected over parametric methods (Brown, 1988; Hall, 2015). Accordingly, the methods of analysis selected were the 2 Related-Samples Wilcoxon Signed Rank Test, Friedman Test, and Kruskal-Wallis Test (Brown, 1988; Mackey & Gass, 2015, Hall, 2015). The 2 Related-Samples Wilcoxon Signed Rank Test was implemented to measure the effect of the digital game tasks on the acquisition of the target vocabulary items and to establish whether or not the participants had learned the target words (Question 1) (Brown, 1988; Mackey & Gass, 2015). This method was applied for two reasons: first, the collected statistical data were rank–ordered data; and second, this method enables comparison of the rank-ordered data means and indicates whether differences between them are significant. Next, the Friedman Test was applied to find differences between participants in their receptive and productive test performances (Question 2). Although there were more than two groups which differed in one independent variable and one-way ANOVA with repeated measure would normally be the most suitable method of analysis in terms of increasing the accuracy of the results, an alternative method, i.e., the Friedman Test, was implemented (Brown, 1988; Hall, 2015). The Friedman Test was expected to reveal which dimension of word knowledge, receptive or productive, was more effectively acquired, after participants had encountered the target words in the digital game tasks. Finally, between-group differences were measured by the Kruskal-Wallis Test to find out which group had outperformed the other two in their post-tests (Question 3). Although ANOVA would normally be the most suitable method of analysis for finding the superior group, the Kruskal–Wallis one-way analysis of variance would, owing to the manner of data collection, be expected to yield more detailed and precise results (Hall, 2015).

4.5.2 Qualitative Data Analysis

Concurrent think-aloud qualitative data were collected to measure the success or failure of each level of involvement load in participants' acquisition of the target vocabulary items in the digital game-based vocabulary acquisition tasks. Content analysis, as described below, was selected as the method of analyzing the concurrent think-aloud data in this study.

> Content analysis is a general term for a number of different strategies used to analyze text […]. It is a systematic coding and categorizing approach used for exploring large amounts of textual information unobtrusively to determine trends and patterns of words used, their frequency, their relationships, and the structures and discourses of communication. (Vaismoradi, Turunen, & Bondas, 2013, p. 400)

Content analysis can, according to the aim of the research, be either quantitative or qualitative (Morgan, 1993; Elo & Kyngäs, 2008). In quantitative content analysis, the researcher preidentifies categories and codes, whereas in qualitative content analysis, the researcher looks for categories, themes, and patterns in the content. Qualitative content analysis is more subjective than quantitative content analysis (Morgan, 1993). In other words, in qualitative content analysis, the researcher's analysis starts from general raw data and ends in a hypothesis or a theory. In quantitative content analysis, in turn, the researcher applies a theory and proceeds from the specific to the general. Elo and Kyngäs (2008, p. 109) offer more practical definitions for qualitative and quantitative content analysis: they define these, respectively, as inductive and deductive approaches. They state that "If there is not enough former knowledge about the phenomenon or if this knowledge is fragmented" the researcher must choose inductive content analysis; otherwise, if "the structure of analysis is operationalized on the basis of previous knowledge and the purpose of the study is theory testing", deductive content analysis must be applied.

The effect of different levels of involvement load on the acquisition of new words in digital game-based vocabulary learning tasks, as studied here, might be considered a new phenomenon. As a researcher, I was seeking to learn more about the distinct impact of each of three levels of involvement load on the vocabulary acquisition. In brief, the aim of this study was to "*detect what has happened*" (Vaismoradi, Turunen, & Bondas, 2013, p. 400) in a context that has been little studied and contains *fragmented* knowledge (Elo & Kyngäs, 2008, p. 109). Therefore, content analysis as the main method of analysis was a logical and an appropriate choice over other methods of qualitative data analysis. Moreover, as a researcher I had to start with raw data that had been collected through concurrent think-aloud protocols. Hence, the techniques of inductive, or qualitative, content analysis were employed as the main method of data analysis in this study.

Elo & Kyngäs (2008) recommend the following steps to a researcher intending to analyze qualitative data by employing inductive content analysis techniques:

Selecting the unit of analysis, such as words, sentences, etc.: Because obtaining knowledge of the processes, strategies and patterns the participants were implementing and following in learning new words at different levels of involvement load was the aim of the present research, the sentence l was considered as the most suitable unit of analysis.

making sense of the data as a whole: The data were transcribed immediately after they were recorded, and they were also reread, listened to and watched several times. Thus, I familiarized myself with the data as a whole the details, settings, participants and events were still fresh in my mind.

open coding: Recurring patterns and themes in the data sheets were coded during reading and rereading.

designing coding sheets: After reviewing the texts, audios, and recorded clips, the patterns that had emerged and their codes were extracted and recorded on separate sheets.

grouping the recurring themes and codes: while watching the video clips, the patterns were also reviewed in order to make sense of the connections between themes and patterns. Next, related patterns were categorized in the same groups.

categorizing the emerged concepts into bigger units: After grouping the patterns, tags were assigned to the units by the most recurring pattern in each group.

abstracting the categories by labeling them: Finally, the patterns were abstracted by labeling them after their concepts in applied linguistics. Each pattern was then grouped into one of two distinct categories, which were abstracted as *universal moves* and *exclusive strategies*. The universal move category contained labels such as information search, negotiation, turn-taking, trial-and-error, and review. The exclusive strategies category contained such labels as input enhancement, planning, inference from the context, hypothesizing, memory search, and feedback request.

reporting the findings, preferably, like a story: To explain the qualitative results, I categorized the findings into two distinct sections, namely universal moves and exclusive strategies. I started with the universal moves. Then, to report the exclusive strategies, I categorized and them based on the relevant group's tag, i.e., A, B or C. In each of the group-specific sections, I present and explain the exclusive strategies; moreover, for all three digital game-based vocabulary acquisition tasks, I describe and illustrate how the emerged categories were connected to each other and formed a group-specific learning approach. Finally, I compared both the universal moves and exclusive strategies between the groups to give the probable reasons why the results of this study partially support previous findings.

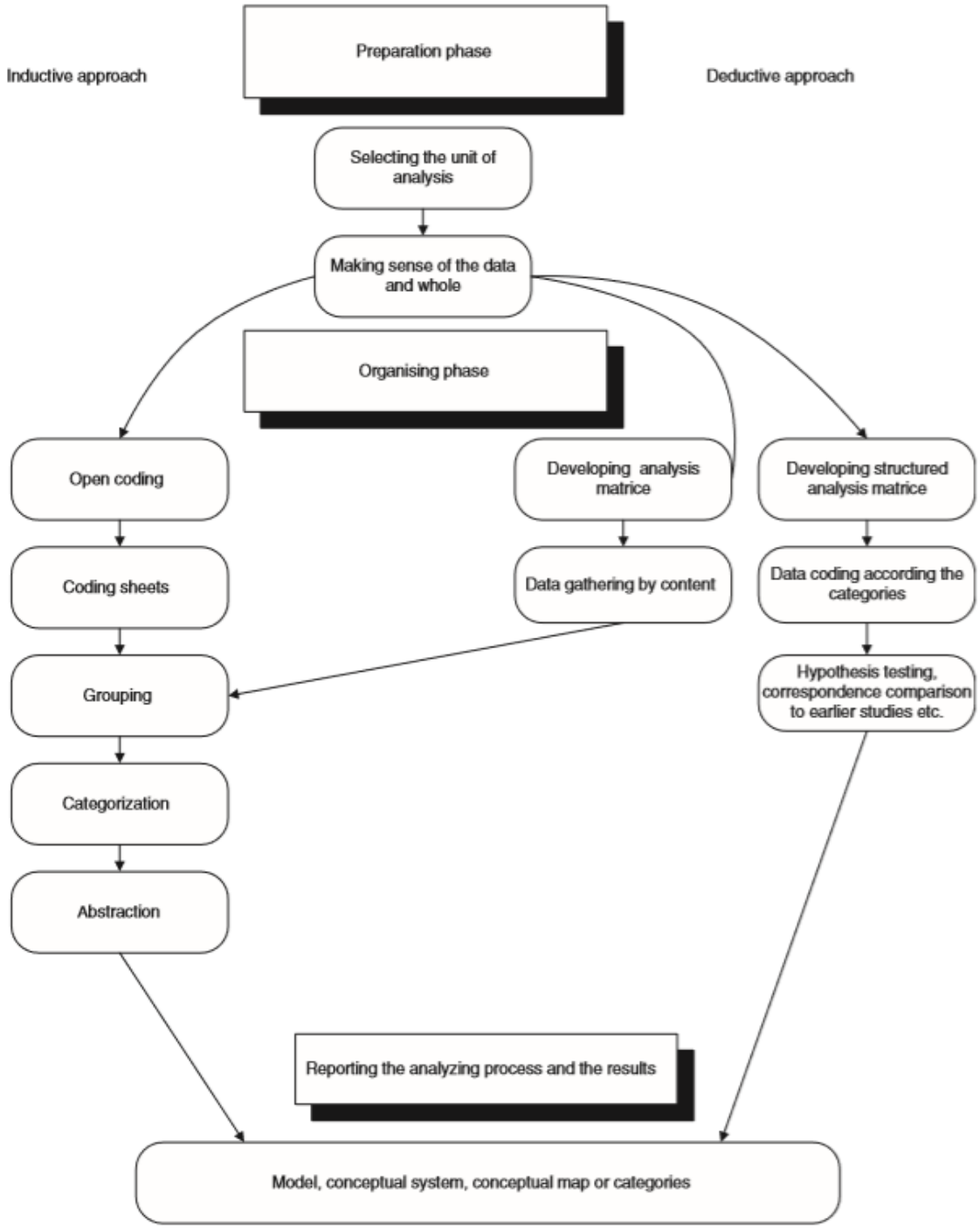


FIGURE 11 Content analysis preparation phases in inductive and deductive approaches (Elo & Kyngäs, 2007, p.110)

5 RESULTS ON VOCABULARY ACQUISITION IN DIGITAL GAME TASKS

In the previous chapter, I described and explained how I had conducted my study and how I had collected the data. In this chapter, I present and describe the findings of the study. In the first section, I describe the quantitative data analysis and report the statistical results on the effect on the acquisition of the target vocabulary items of the three digital game tasks inducing three different levels of involvement load. In so doing, I have broken down the main question into four smaller questions, each of which required its own distinct statistical analysis. In the second section, I describe the results of the qualitative data analysis. I describe and explain the learning processes and patterns applied and generated by the participants in their different groups when they were engaging with their digital game tasks. Moreover, I offer tentative explanations for the success or failure of their chosen learning paths in acquiring the target vocabulary items. Finally, I describe the models of vocabulary acquisition that emerged owing to the different levels of involvement load in each task.

5.1 Quantitative Results of the Achievement Tests

5.1.1 Descriptive Statistics

This study comprised 30 participants. Their vocabulary knowledge was tested by two pre-tests and two post-tests before and after the data collection procedure. The tests measured participants' receptive (recognition/recall) and productive (recognition/recall) knowledge of the target vocabulary items before and after playing the digital game. Table 16 presents the descriptive statistics on the participants' performance in both the receptive and productive tests in the pre- and post- phases of test administration. The table includes means and standard deviations, showing the differences between participants in their vocabulary acquisition. Three different tasks inducing three different levels of involvement

load were performed. The three levels were implemented by making slight variations in the structure of the tasks, as described in the previous chapter. It seems that the slight differences in the tasks led to wide variation among the participants in their task performance and vocabulary acquisition. As the rough data show, the participants in group C achieved the highest means in both the productive (17.4) and receptive tests (16.9). The same was true of the sub-tests, i.e. recognition and recall. Group C's tasks induced the highest level of involvement load, These results were predictable based on the literature (Laufer & Hulstijn, 2001; Jing & Jianbin, 2009). In contrast, and surprisingly, the group B participants, with the moderate level of task-inducing involvement load, scored lowest in the productive post-tests (12.00). Group B's result was surprising as it contradicted previous findings, according to which a moderate Level of involvement load would be expected to lead to second-best performance in a vocabulary acquisition task also performed at a higher and lower level of involvement load (Laufer & Hulstijn, 2001). However, these inferences and interpretations are based solely on the raw data, shown in table 16, and to be either confirmed or rejected would require deeper analysis.

The raw data were analyzed by nonparametric methods of data analysis, as these are the most reliable for comparing rank-ordered data and mean ranks in small data sets (Brown, 1988; Mackey & Gass, 2015, Hall, 2015). Moreover, the outcomes of non-parametric data analysis are significantly more accurate in research with a small number of participants. Therefore, to increase the accuracy of the analysis and the quality of the results, the main question was broken down into three smaller questions. Details of the analysis and the results of the non-parametric analyses of the data for each of the three questions are described below in separate sections.

TABLE 16 Descriptive statistics of receptive and productive pre- / post-tests

Group		Pre-Test Productive Recall	Pre-Test Productive Recognition	**Pre-Test Productive**	Post-Test Productive Recall	Post-Test Productive Recognition	**Post-Test Productive**	Pre-Test Receptive Recall	Pre-Test Receptive Recognition	**Pre-Test Receptive**	Post-Test Receptive Recall	Post-Test Receptive Recognition	**Post-Test Receptive**	**Post - Pre* Productive**	**Post - Pre* Receptive**
Group C	N	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	Mean	2.4000	2.1000	4.5000	8.2000	9.2000	17.4000	2.1000	2.8000	4.9000	8.0000	8.9000	16.900	12.900	12.00
	Std. Deviation	2.79682	2.42441	4.52769	2.20101	.78881	2.91357	2.6436	1.87380	3.81372	2.4494	1.59513	3.9001	5.3218	5.557
Group B	N	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	Mean	2.9000	4.0000	6.9000	5.6000	6.4000	12.0000	3.4000	4.2000	7.4000	5.6000	6.6000	12.200	5.1000	4.800
	Std. Deviation	2.76687	2.05480	4.33205	3.33999	2.95146	5.96285	3.2386	2.25093	5.52167	3.0258	2.75681	5.6725	2.7264	2.859
Group A	N	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	Mean	1.1000	4.2000	5.3000	5.1000	7.6000	12.7000	1.3000	4.0000	5.3000	4.1000	6.3000	10.400	7.4000	5.400
	Std. Deviation	1.10050	2.20101	2.79086	2.13177	1.83787	12.7000	1.3374	2.21108	3.30151	1.7288	1.82878	3.1693	2.9888	3.169
Total	N	30	30	30	30	30	30	30	30	30	30	30	30	30	30
	Mean	2.1333	3.4333	5.5667	6.3000	7.7333	14.0333	2.2667	3.6667	5.8667	5.9000	7.2667	13.166	8.4667	7.400
	Std. Deviation	2.40306	2.35889	3.95390	2.87858	2.30342	4.75962	2.5988	2.13886	4.31304	2.8809	2.36254	5.0588	4.9947	5.123

** the (-) sign is subtraction*

5.1.2 Inferential Statistics

5.1.2.1 Question 1: What is the effect of the digital game, in different levels of involvement load, on the acquisition of target vocabulary items?

In the first phase of the study, it was imperative to find out whether or not the digital game tasks were effective enough to enhance acquisition of the target vocabulary items. To do this, the effect of the digital game must be measured twice as the mental constructs of receptive and productive knowledge of vocabulary are considered separate and distinct (Nation, 2001). Hence, the participants' word knowledge was evaluated by both productive and receptive tests and the mean ranks of participants' performance were compared separately for each of the two types of test. This meant answering two questions: First,

5.1.2.1.1 Does playing the digital game make a significant difference in the participants' performance, in their productive tests, before and after playing the digital game?

To answer this question, participants' productive knowledge of the target vocabulary items, both before and after playing the digital game, was compared to find whether or not the digital game had any significant effect on the participants' minds and assisted them in developing productive knowledge of the target vocabulary items. The outcomes of the tests were ten mean ranks, indicating participants' performance in both test administration phases.

After calculating the mean ranks of the both the productive pre –and post – tests, the 2 Related-Samples Wilcoxon Signed Rank Test in SSPS was applied. This was done for two reasons: first, the collected statistical data were rank-ordered data; and second, this method enables comparison of the means of rank-ordered data, and indicates if the differences between them are significant. After applying the method, the outcome of the analysis was as follows (table 17).

TABLE 17 Nonparametric comparison of PRODUCTIVE pre-tests and post-tests

		N	Mean Rank	Sum of Ranks	Z	Asymp. Sig. (2-tailed)
Post - Productive (Recognition) Pre – Productive (Recognition)	Negative Ranks	1[a]	3.50	3.50	4.724	.000
	Positive Ranks	29[b]	15.91	461.50		
	Ties	0[c]				
	Total	30				
Post – Productive (Recall) Pre – Productive (Recall)	Negative Ranks	0[d]	.00	.00	4.799	.000
	Positive Ranks	30[e]	15.50	465.00		
	Ties	0[f]				
	Total	30				
Post - Test Productive Pre - Test Productive	Negative Ranks	0[g]	.00	.00	4.709	.000
	Positive Ranks	29[h]	15.00	435.00		
	Ties	1[i]				
	Total	30				

Table 17 shows that 29 participants scored higher in their productive post-tests and that their overall mean rank was 15.00. However, one participant had the same score in both the productive pre-test and productive post- test. The z-score of 4.709, at $p \leq 0.05$ level[4], indicates a significant difference between the mean ranks of the productive pre- and post-tests.

Moreover, statistical results for the test sub-sections, i.e. recognition and recall, indicated that 29 participants scored higher in the productive recognition post-test and that their mean rank in this sub-section was 15.91, which was also higher than that of the participant who scored better in the productive recognition pre- than post-test , i.e. 3.50. The z-score of 4.742, at the significance level of $p \leq 0.05$, shows a significant difference between the mean ranks of the productive recognition pre- and post-test after playing the digital game.

Finally, the results of the recall sub section of the productive test revealed that all 30 participants scored significantly higher in their productive recall post- than pre-test. Their mean rank in the productive recall post-test was 15.50 and the z score 4.799 at the significance level of $p \leq 0.05$.

The above results indicate that the digital game enhanced the participants' acquisition of productive knowledge of the target vocabulary items. This finding is important as previous studies on vocabulary acquisition have found that the acquisition of productive knowledge and recall of target words is more challenging than the acquisition of receptive knowledge and word recognition (Sökmen, 1997; Nation, 2001; Schmitt, 2007, 2008). However, the present results indicate effective acquisition of productive and recall word knowledge for the

4 The level of significance, i.e the value that determines if differences between compared mean ranks are significant. Any score higher than 0.05 in Asymp. Sig. (2-tailed) indicates that the difference between the compared mean ranks is not significant.

target vocabulary items in this study. Another interpretation of this finding is that the integration of digital games into language classrooms can be very supportive in vocabulary acquisition, even in such difficult and challenging aspects as acquiring productive/recall knowledge of target words.

5.1.2.1.2 Does playing the digital game make a significant difference in the participants' performance, in their receptive tests, before and after playing the digital game?

As in the previous analysis for the productive tests, the mean ranks of the both receptive pre–tests and receptive post-tests were compared using the 2 Related-Samples Wilcoxon Signed Rank Test by SPSS.

TABLE 18 Nonparametric comparison of RECEPTIVE pre-tests and post-tests

		N	Mean Rank	Sum of Ranks	Z	Asymp. Sig. (2-tailed)
Post - Receptive (Recognition) Pre – Receptive (Recognition)	Negative Ranks	0[a]	.00	.00	4.635	.000
	Positive Ranks	28[b]	14.50	406.00		
	Ties	2[c]				
	Total	30				
Post – Receptive (Recall) Pre – Receptive (Recall)	Negative Ranks	1[d]	4.50	4.50	4.624	.000
	Positive Ranks	28[e]	15.38	430.50		
	Ties	1[f]				
	Total	30				
Post - Test Receptive Pre - Test Receptive	Negative Ranks	0[g]	.00	.00	4.790	.000
	Positive Ranks	30[h]	15.50	465.00		
	Ties	0[i]				
	Total	30				

The results showed that all participants (30) scored higher in their receptive post-tests and that their overall mean rank was 15.50. The z score of 4.790 at the significance level of ($p \leq 0.05$) indicated a significant difference between the mean ranks of the receptive pre- and post-tests (Table 18). That is, the results showed that participants performed the receptive post-tests better than the receptive pre-tests.

The statistical results for the sub-sections of this test indicated that 28 participants scored higher in the receptive recognition, post- than pre-test. Although 2 participants' scores were the same for both the receptive recognition) pre- and post-tests at the significance level of $p \leq 0.05$), the z score of 4.635 indicated a significant difference between the participants' mean ranks in both receptive recognition pre- and post-tests after playing the digital game.

The results show that 28 participants scored higher in the recall sub-section of the receptive post-tests. However, 1 participant performed equally in both the

receptive recall pre- and post-tests and 1 other participant performed better in the receptive recall pre-test. Hence, the mean ranks of the 28 participants who performed better in the receptive recall post-test, i.e. (15.38), was higher than the mean rank of the participant who performed better in the receptive recall pre-test. The z score of 4.626 at the significance level of $p \leq 0.05$ indicated a significant difference between the mean ranks of the participants' performance in receptive recall pre- post-tests. Thus, as with productive knowledge, the digital game also assisted the participants in acquiring receptive knowledge of the target vocabulary items. Moreover, it also enhanced acquisition of the recognition and recall aspects of the target vocabulary items.

The outcomes of the two analyses were as expected and supported previous findings (Yip & Kwan, 2006; Fotouhi-Ghazvini, Earnshaw, Robinson, & Excell, 2009; Laveborn, 2009; Hung, 2011; Vahdat & Rasti Behbahani, 2013; Chen, Tseng, Hsiao, 2018). The results indicate that the digital games were successful in boosting vocabulary acquisition in both the receptive and productive dimensions of word knowledge. Moreover, the results added the novel finding that the recall and recognition scopes of word knowledge might also be enhanced and supported through a digital game-based vocabulary acquisition task. In addition, the results emphasized the positive role of the digital game in vocabulary acquisition irrespective of the manner and structure of the task, in this instance levels of involvement load. Thus, I concluded that vocabulary acquisition was enhanced by playing the digital game; hence, the later analysis showed that the data measured a mental construct that had been established by playing the digital game, not by any unexpected factors. Thus, the answer to the first question of this study is that the digital game probably contributed considerably to the acquisition of the target vocabulary items.

5.1.2.2 Question 2: Which dimension and scope of word knowledge, either receptive (recall/recognition) or productive (recall/recognition), are acquired significantly better after completing digital game tasks in different levels of involvement load?

The first phase of the analysis showed that that playing the digital game had assisted the participants to acquire the target vocabulary items both receptively and productively. The aim of the second phase of the analysis was to find out which *dimensions* and *scopes* of word knowledge had been influenced the most when the digital game tasks were performed at different levels of involvement load Because the analysis involved more than two groups that differed in one independent variable, one-way ANOVA with repeated measure would normally be considered the most suitable method of analysis. However, since, as already mentioned, the data had been gathered from only a small population, an alternative method that would increase the accuracy of results was selected (Brown, 1988; Mackey & Gass, 2015). Thus, the Friedman test SPSS was used to compare the mean ranks of both the receptive and productive tests (Hall, 2015). Moreover, the Friedman test was the best choice for comparing the mean ranks

because it applies to rank-ordered data and compares the mean ranks of more than two groups that differ in one independent variable. The outcomes of the analysis are presented in table 19.

TABLE 19 Comparison of mean rank of productive and receptive tests

	Mean Rank	N	Chi-Square	df	sig
Post.test Productive	17.33	30	303.022	17	0.000
Post.test Receptive	17.03				
Post.test Productive (Recognition)	13.13				
Post.test Receptive (Recognition)	12.70				
*Post-Pre Productive	12.67				
*Post-Pre Receptive	11.37				
Post.test Productive (Recall)	10.45				
Pre.test Receptive	10.40				
Pre.test Productive	10.17				
Post.test Receptive (Recall)	9.72				
Pre.test Receptive (Recognition)	7.10				
*Post-Pre Productive (Recognition)	6.88				
*Post-Pre Productive (Recall)	6.68				
Pre. Productive (Recognition)	6.48				
*Post-Pre Receptive (Recall)	5.60				
*Post-Pre Receptive (Recognition)	5.53				
Pre.test Receptive (Recall)	4.00				
Pre.test Productive (Recall)	3.75				
* the (-) sign is subtraction					

The results revealed that the mean rank of the participants' performance was highest (17.33) in the productive post-test and lowest (3.75) in the productive recall pre-test. The chi-square value (303.022) at the significance level of $p \leq 0.05$ indicated a significant difference between the mean ranks of the participants' performance in the productive and receptive tests and their sub sections, recall and recognition. In other words, the results showed that the participants had performed better in the productive post–test; that is, their productive knowledge of the target vocabulary items improved more than their receptive knowledge of the same items after playing the digital game.

The results for the recognition and recall scopes of the target words showed that participants had acquired recognition better than recall. According to table 19, the participants had performed best in the productive recognition post-test (13.13), followed by the receptive recognition post-test (12.70). Participants' performance in the recall section of the productive post-test (10.45) was also significantly better than recall in the receptive post-test (9.72). The chi-square value (303.022) for the differences between the recall and recognition sections of the tests was also significant. It can be concluded that the recognition scope of the target vocabulary items was more effectively acquired than recall scope of the same vocabulary items in the digital game-based vocabulary acquisition tasks.

The findings, in this section, support the those of Jasso (2012) and Sundqvist (2019). Both authors also found that a commercial digital game better assisted the

acquisition of productive than receptive knowledge of target words. These findings in favor of productive over receptive knowledge acquisition do not support most previous reports. Several studies have reported that vocabulary items were primarily acquired receptively (Morton, 1977 as cited in Barcroft, 2004; Meara, 1997; Nation, 1990, 2001; Schmitt, 2010a). Moreover, in Meara's model of word learning, the lexicon develops from receptive to productive (Schmitt, 2010a). The findings contradict those of earlier studies and bring to mind Melka's views on the receptive and productive knowledge of words (1997, as cited in Barcroft, 2004). Melka argued against the notion of two distinct types of vocabulary knowledges, i.e. receptive and productive. Instead, he suggested that the receptive/productive distinction might only exist on the level of testing rather than in learners' minds. Thus, it can be inferred that, in the present study, the digital game had a stronger influence on performance in the productive tests.

5.1.2.3 Question 3: Does interacting with the digital game tasks, in different levels of involvement load, make significant differences in vocabulary acquisition?

Up to this point, it has been shown that the digital game assisted the participants in learning the target words. Moreover, the digital game clearly better supported their acquisition of productive than receptive knowledge of the target words. In the last phase of the analysis, the effect of different levels of involvement load on the acquisition of the target vocabulary items while playing the digital game was measured. In other words, I investigated whether working with a high, moderate or low level of involvement load produced superior results in any of the productive and receptive post-tests.

Thus, the main aim of the analysis was to find out if the three groups, each with a different level of involvement load, would differ in their vocabulary acquisition. As the amount off collected data was small, the mean ranks of the groups' performance in both the receptive and productive tests were compared. Due to the nature of the data and its ability to deliver more detailed and precise results, the Kruskal-Wallis one-way analysis of variance was favored over ANOVA, (Hall, 2015).

TABLE 20 Between-group differences in productive and receptive post-test outcomes after playing the video game

Tests	Groups	Mean Rank	N	Chi-Square	df	sig
(post-*pre) Productive	Group A	14.35	10	10.687	2	0.005
	Group B	9.75				
	Group C	22.40				
(post-*pre) Receptive	Group A	12.90	10	9.240	2	0.010
	Group B	11.30				
	Group C	22.30				
(post-*pre) Productive (Recognition)	Group A	13.00	10	12.892	2	0.002
	Group B	10.10				
	Group C	23.40				

(post-*pre) Productive (Recall)	Group A	16.05	10	5.031	2	0.081
	Group B	10.90				
	Group C	19.55				
(post-*pre) Receptive (Recognition)	Group A	11.45	10	10.838	2	0.004
	Group B	12.15				
	Group C	22.90				
(post-*pre) Receptive (Recall)	Group A	13.50	10	7.399	2	0.025
	Group B	11.50				
	Group C	21.50				
* the (-) sign is subtraction						

The results showed that, in the productive post-test, group C, which experienced learning new words with the highest level of involvement load when playing the digital game, scored highest (mean rank = 22.40) and that Group B, with the moderate level of involvement load, scored lowest (mean rank = 9.75) (Figure 12). The chi-square value (10.687) with 2 degrees of freedom at the significance level of $p \leq 0.05$ indicated a significant difference between the groups.

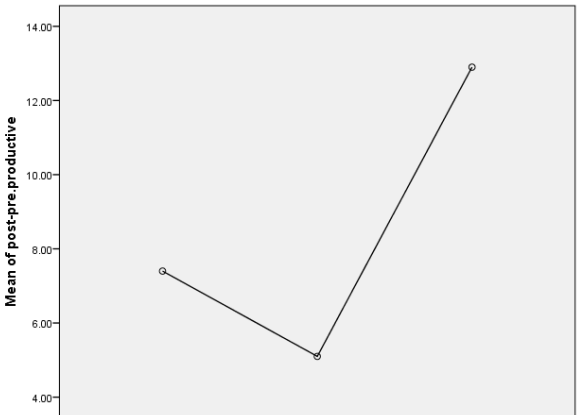


FIGURE 12 Comparing the means of the post-test (productive)

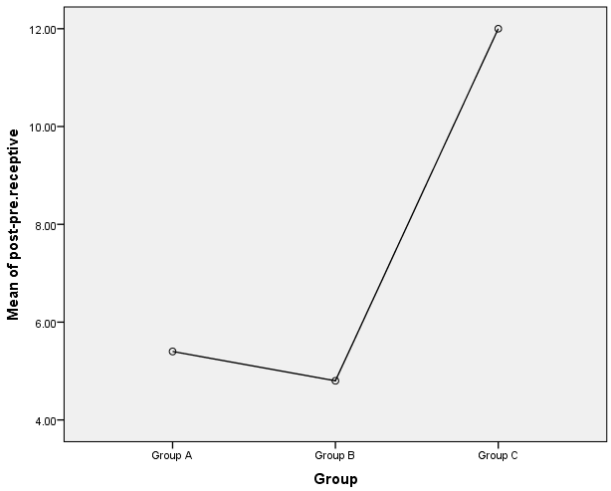


FIGURE 13 Comparing the means of the post-test (receptive)

Furthermore, in the receptive post-test, group C again showed the highest score (mean rank = 22.30 and group B, with the moderate involvement load, the lowest (mean rank = 11.30) (Figure 13). Moreover, at the significance level of $p \leq 0.05$ with 2 degrees of freedom, the chi-square value (9.240) indicated a significant difference between the groups.

In the recall and recognition scopes of knowledge of the target words, group C showed the highest score and mean rank (23.40) in the productive recognition post-test, followed by group A (mean rank = 13.00) and group B (mean rank = 10.10) (Figure. 14). At the significance level of $p \leq 0.05$ with a 2 degrees of freedom, the chi-square value (12.892) indicated a significant difference between the three groups.

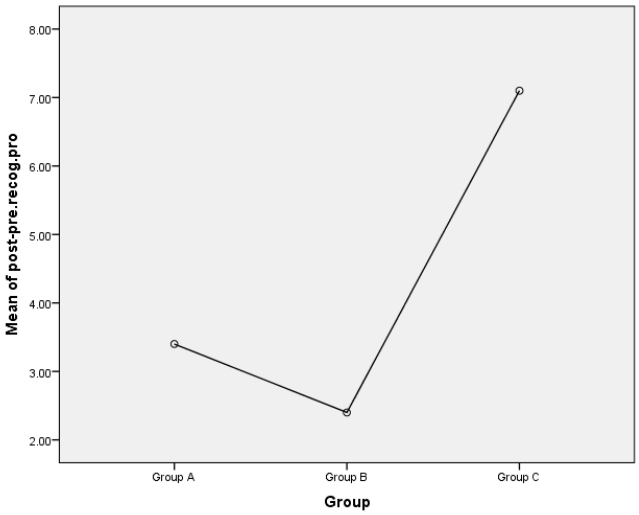


FIGURE 14 Comparing the means of the post-test productive (recognition)

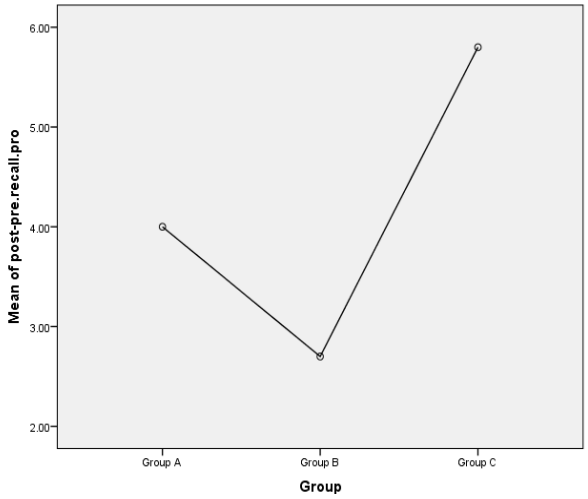


FIGURE 15 Comparing the means of the post-test productive (recall)

In the productive recall post-test, group C again showed the highest score and highest mean rank (19.55), followed by group A (mean rank=16.05 and group B (mean rank = 10.90) (Figure. 15). However, in this comparison, the chi-square value (5.031) at the significant level of $p \leq 0.05$ with 2 degrees of freedom indicated that the difference between the groups in the productive recall post-test was due to other unanticipated factors, such as context, mind processes, strategies etc. that will be discussed later in this chapter, and that the levels of involvement load played no significant role in this regard.

In the receptive recognition post-test, group C was again showed the highest score (mean rank = 22.90), followed by group B (mean rank = 12.15) and group A (mean rank = 11.45) (figure 16). Moreover, at the significance level of $p \leq 0.05$ with 2 degrees of freedom, the chi-square value (10.838) indicated a significant difference between the groups.

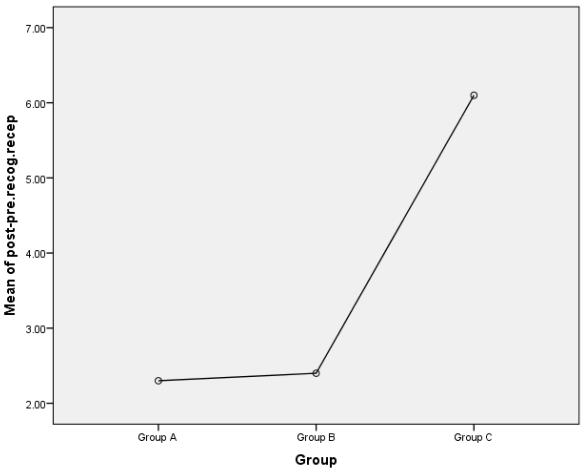


FIGURE 16 Comparing the means of the post-test receptive (recognition)

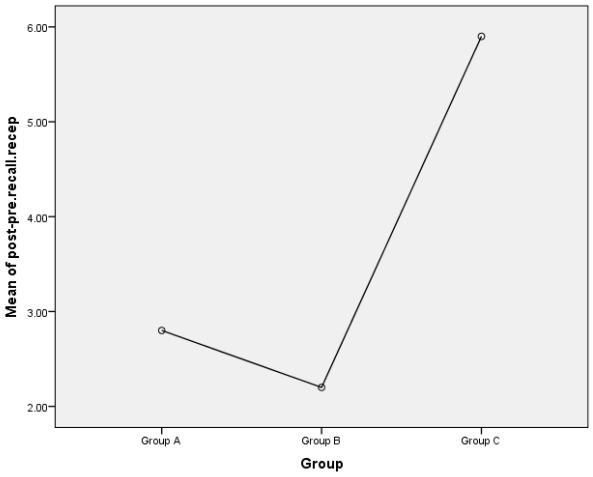


FIGURE 17 Comparing the means of the post-test receptive (recall)

Finally, in the receptive recall post-test, as in the previous test, group C outperformed the other groups (mean rank = 21.50), followed by group (13.50) and group B (11.50) (Figure. 17). The Chi-square value at the significance level of $p \leq 0.05$) with 2 degrees of freedom indicated a significant difference between the groups.

These results showing that group C outperformed the other two groups in both the receptive and productive post-tests and ranked the highest in both suggest that the group C participants' high scores were due to their engagement with the digital game task inducing a high level of involvement load. The statistical analyses also showed significant between-group differences in performance at the different levels of involvement load. Therefore, the superior performance of group C, in acquiring the target vocabulary items, was probably mostly due to the nature of the digital game task that had induced a high level of involvement load. Furthermore, the results revealed that the Group B participants, who had performed the digital game task with a moderate level of involvement load, showed the lowest scores in both the receptive and productive post-tests. This result was counter to expectations as, according to the involvement load hypothesis (Laufer & Hulstijn, 2001), they should have shown the second-best performance in the post-tests. However, group A, which had the lowest level of involvement load, showed the second-best performance in the post-tests. Thus, although these findings point to a possible positive effect of level of involvement

load on the acquisition of the target vocabulary items in digital game tasks, they also signal the possible existence of other contributory factors that might induce unanticipated negative or positive influences on the expected outcome of such tasks. I discuss these factors further in this chapter in connection with the results of the qualitative analysis.

In previous studies, researchers have suggested that level of involvement load is an important psychological determinant in vocabulary acquisition (Laufer & Hulstijn, 2001; Tsubaki, 2007). The results of the present analysis were in line with the majority of empirical studies in the literature (Jing & Jianbin, 2009; Kim, 2011; Huang, Wilson, Eslami, 2012; Xie, Zou, Wang, Wong, 2017; Zou, 2017). Moreover, the performance of group B supported the findings of a study in which it was concluded that "the operationalization of the levels of processing [...] needs reconsideration" and "the involvement index may well not function independently of the task type for vocabulary instruction. That is, the processing load brought to bear by task type may well affect word retention" (Yaqubi, Rayati, & Gorji, 2012, p. 104, 161). A more in-depth discussion of the results and probable contributory factors is discussed in detail in chapter 6.

5.2 Qualitative Results

In this section, I report the findings of the inductive content analysis of the collected concurrent think-aloud data. With respect to the controversial quantitative findings of this study pertaining to the performance of group B in the post-tests, the data gathered on the participants' mental online-processing were explored to find out how each of the three levels of involvement load might have affected their performance during their engagement with the digital game tasks. The participants' learning patterns might reveal factors that either enhanced or hindered their vocabulary acquisition via the digital game task.

To explain the qualitative results, I have categorized the findings into two sections, namely, universal moves, that is, moves that were not group-specific and followed by all participants regardless of their task type, and exclusive strategies, that is, strategies that were group-specific and hence task-related. I start by discussing the universal moves. I then categorize and describe the exclusive strategies used by each of the three groups. Moreover, I describe and illustrate how the categories that emerged, i.e. universal moves and exclusive strategies, were connected to each other and formed a group-specific problem-solving model, or a learning approach, for each of the three digital game-based vocabulary acquisition tasks. Finally, I compare the universal moves and exclusive strategies to identify the probable reasons why the results of this study partially support previous findings in the literature.

5.2.1 Universal Moves

The data analysis revealed that all the participants repeated five moves in completing the digital game tasks. These moves were performed by the participants in almost every phase of the task. The five universal moves were labelled 1) information search, 2) negotiation, 3) turn-taking, 4) trial-and-error, and 5) review. Among the categories that emerged, those were the most iterated moves. As mentioned earlier, I named the categories based on their internal attributes and similarities to such concepts in Applied Linguistics as *input enhancement, inferencing,* and *feedback requests*. Below, I explain them in detail. Their frequency of occurrence is shown in table 21.

TABLE 21 Frequencies of the Categories of Universal Moves

	Group A	Group B	Group C
Information Search	156	106	132
Negotiation	81	150	59
Turn-taking	38	61	43
Trial-and-Error	40	49	24
Review	56	18	59

5.2.1.1 Information Search

To complete any task requires explanations and instructions (Ellis, 2017). If the digital game task is considered as a combination of relevant tasks, participants need instructions, explanations and information to develop strategies and plans for completing each task. Although the present participants were given a general instruction on how to complete the digital-game task, the only detailed sources of information they had were the game guide and the digital game; thus, to complete the in-game tasks and to progress in the game, they had no choice but to search for information in both the game guide and the game itself. Hence, they checked both the game guide and the game for information about the target words or to monitor their current situation. These instances when the participants referred to the text or the game for the purpose of either moving on to the next task or obtaining more information about their current task were coded as information search

Shirin	Shadi
خوب میگه چی؟ Click on the BROKEN DOOR	
Click on the BROKEN DOOR	
این چیه ، 3/1، 3/2؟	
	یه سوم دو سوم
	شاید باید بهش وصلش کنی!
I could… ها ، چسب	
	نه چسب کناریشه خوب ، حالا باید بری بیرون همونجا
اه {در حال تست کردن}	
	Now I can repair …

Shirin	Shadi
So, what does it read? Click on the BROKEN DOOR	
Click on the BROKEN DOOR	
What is this 3/2 ،3/1 ?	
	One third, two third?
	Maybe, you should attach them
I could… **{information from the game}** **Aha, glue**	
	No, glue is the next one Ok, now you should return back there
Eh… {testing the glue}	
	Now I can repair… {game info}

Ali 2	Alireza
ولی گفت گلو رو کجا پیدا کنین؟	
یه بار دیگه این صفحه رو نگاه کن	
Click on the LADDER منظورش این درای آویزون بود	باید یه جایی گلو داشته باشیم
click on the glue	
{در حال کلیک و جستجو}	برو اون درو باز کن

Ali 2	Alireza
But it noted where to find the glue	
Look at this page once more	
Click on the LADDER Does ladder mean curtain door?	We must have glue somewhere
click on the glue	
{**In-game information search**}	Open that door

EXAMPLE 1 An excerpt showing examples of information search

As it is evident in the two examples, the participants referred to both the game and the game guide in order to understand how to find and apply the "glue" in order to complete one of the game tasks.

5.2.1.2 Negotiation

As mentioned earlier, the digital game task was completed in pairs. After finding the correct information, the participants had to develop either strategies or plans for solving the in-game tasks. Because the digital-game tasks were completed in pairs, the participants negotiated their choice of strategy and implemented it. Thus, their joint involvement and negotiation in

finding the best solution to the problems in the tasks was inevitable. The proposals of one participant were either accepted or rejected by the other, who might then offer alternative solution. The purposes of negotiations were either explanations or justifications. The participants negotiated over words and proposed solutions. In the following example, the participants are negotiating over the target word "ladder". They were trying to find the Persian definition of the word but in this example were unsuccessful.

Tara	Negar
Click on the LADDER	
لدر چی بود؟	من میدونستم شبیه چی میشد
من لتر رو می دونم که شبیه چرم میشه	من لدر رو اصلا یادم نمیاد
این چیه؟	آها این برگه هاس ، پیدا شد
نمیدونم فعلا اینو بردار	الان لدر چی میشد؟
Go into the house. The door will close fast	شاید لدر منظورش این جعبه کمک های اولیه س

Tara	Negar
Click on the LADDER	
What was ladder?	I can imagine the shape
Wasn't it "leather"?	I don't remember "ladder"
What is this?	Aha, these leaves, found it!
I don't know, collect this	So, what is ladder?
{she reads the game guide}	Maybe, by ladder it means first-aid kit

EXAMPLE 2 An excerpt showing an example of negotiation over target words

In the next example, the participants negotiate a plan. One wants to ensure that they understand of the instruction for completing the task.

Minoo	Mähdieh
	Click on the ladder; then, Click on the 3 planks; Click on the 6 nails in your bag
مطمعنی باید بکنیم؟	ما باید سه تا داشته باشیم ، اون سه تا رو از کجا بیاریم؟
	نگاه گفته پلنکمون سه تا باشه
میدونم	
	Click on the LADDER in your bag
خوب میگه وقتی که سه تا شد کلیک کن	{کلیک ... کلیک}

Minoo	Mähdieh
	Click on the ladder; then, Click on the 3 planks; Click on the 6 nails in your bag
Are you sure on this?	We must have three of them, where should we find?
	Look, it says we must have three planks
I know	
	Click on the LADDER in your bag
Maybe, it says click when you have all three	{Click … Click}

EXAMPLE 3 An excerpt showing an example of negotiation over planning future actions

5.2.1.3 Turn-Taking

After finding the correct information, the participants negotiated and agreed on their strategies, solutions, and ideas for solving problems. They evaluated their new ideas and strategies by taking turns. Turn-taking was done either by controlling the game with mouse clicks or by interacting with the game guide text. Because there was only one device for each pair, they needed to take turns to test their ideas and strategies. In the following example, the group B participants take turns by swapping possession of the mouse to test their constantly evolving ideas for solving a problem.

Ali	Moein
برو برو برو یه لحظه بده {عوض کردن کنترل کننده}	
{کلیک ... کلیک در حال تست کردن}	
هیچ نمیشه	چی برداشتی؟
Click on the HAMMER	خوب باید رو چکش بزنی
ها میدونی میگه چی کار کن؟{عوض کردن کنترل کننده}{کلیک ... کلیک} در حال تست کردن	

Ali	Moein
Go, go, go Give it to me {exchanging the mouse}	
{Click… Testing his idea}	
Nothing It didn't work	What did you take?
Click on the HAMMER	So, you must click on the hammer
Aaha, do you know what it asks us to do?{Exchanging the mouse}{Clicks for testing his idea}	

EXAMPLE 4 An excerpt showing an example of turn-taking in controlling the game

5.2.1.4 Trial-and-Error

When none of the plans or hypotheses had succeeded, the final option for the participants in each group was trial-and-error. In this move, the participants benefited from the immediate feedback (Kiili, 2005) feature of digital games. As previously mentioned, digital games provide gamers with immediate feedback if their actions have not previously been programmed into the game. For example, in the scene in this game where the participants must find a specific object, such as a *door sign*, in order to progress to the next level or task, the gamers' clicks on every other object is ineffective and the game does not progress. In this case, clicking and lack of progress constitute immediate feedback for the gamers. Thus, the gamers know that their actions were wrong, and that they need to continue trying and clicking on other objects in the scene. They repeat the act of clicking until they click on the correct object and are allowed to progress to the next task. I labelled this move trial-and-error. This move was often implicit and found by watching the participants' in-game gameplay videos. The following excerpt clarifies the notion of the trial-and-error move. In the following excerpt, Shirin and Shadi were trying to find the door *sign* and all their attempts so far had failed. Consequently, they relied on trial-and-error.

Shirin	Shadi
click on the door **sign**	علامت در کجاس؟ علامت در اینه دیگه
{کلیک ، کلیک...}	علاامتـه در
	این نیس علامت در؟
نچ	

Shirin	Shadi
click on the door **sign**	Where is the door sign? Isn't that the door sign?
{Clicking frequently…}	doooor siiiign (thinking aloud)
	Isn't that the door sign?
nope	

EXAMPLE 5 An excerpt showing an example of a trial-and-error move

Moreover, in the exit interview, when asked "Do you use any tricks, when playing a video game, to help you with unknown words? What about in this game?" some of the participants affirmed that trial-and-error helped them to solve problems in their digital game tasks.

Do you use any tricks, when playing a video game, to help you with unknown words? What about this game?	
Ali2	**Alireza**
بعضی کلمه هارو که نخونده بودیم واصلا بلد نبودیم با آزمون و خطا پیدا می کردیم	آزمون و خطا
We found some words that we didn't read or know by trial and error.	*Trial and error*

EXAMPLE 6 Participants mentioned using trial-and-error

5.2.1.5 Review

Finally, participants reviewed previously solved problems regularly to obtain a better understanding of their current situation and reorganize their minds. By review, I mean they returned to completed actions either in the game guide or in the game. The extent and scope of reviewing was mostly limited to the problem before the last one. For instance, in the following excerpt, the participants were reviewing the number and the places of the nails that they needed to find.

Ali 2	Alireza
Go back خوب گفت down.	
	بعد Click the HAMMER in your bag, click on the **portrait**
	پورتریت چی بود؟
عکس	

Ali 2	Alireza
Well, it said Go back down	
	Then Click the HAMMER in your bag, click on the **portrait**
	What was portrait?
{Persian definition}	

Behind that, Click also on the PLANK 1/3 and ISAAC'S REEL میگه پشتش	ها
خوب این میخواشم کندیم	
	2تا ا 6 تاس ، باید بازم بکنی
پشتش... ، آیزاک... ، اینو هم پیدا کردیم.	

Behind that, Click also on the PLANK 1/3 and ISAAC'S REEL It mentions behind	Aha
Ok, we removed its nails either	
	2/6, you need more
Behind that…, Isaac…., We found it as well	

EXAMPLE 7 An excerpt showing an example of reviewing

All five moves, i.e. information search, negotiation, turn-taking, trial-and-error, and review, were universal moves made by all the participants in every group. Although universal, their frequency of occurrence differed depending on the group's task. In a later section of this chapter, the probable reasons for the frequency of these moves and why they varied across the groups is discussed.

5.2.2 Exclusive Strategies

To induce three different levels of involvement load, three tasks with different structures were designed. Although each group's tasks were different and had different levels of involvement load, universal moves were evident and frequent in the participants' think-aloud data regardless of the differences in the task structures. However, the differences in the task structures led each group to employ strategies that were task-/group–specific. These strategies were specifically for overcoming either target or new word-related challenges such as finding, recalling, and guessing their definitions. These strategies were exclusive to each group, that is, each task, with its different level of involvement load. Below, these strategies are categorized and explained.

5.2.2.1 Group A (Lowest level of involvement load)

Scrutiny of the group A participants' online processes showed that they employed a specific strategy, known as the **input enhancement** strategy, when they encountered the target words. This strategy was exclusively used by this group. In using the strategy, they either voiced the target words that they encountered, or they enunciated the Persian definition given in the English language game guide. For example, in the following excerpt, Mähdieh, who was reading the game guide in English, enhanced the input by enunciating the Persian definition instead of *closet,* the English target word. Minoo, who was listening, repeated the definition.

Minoo	Mähideh
گنجه	Click on the HAMMER in your bag; click on the red گنجه

Minoo	Mähideh
{Persian definition}	Click on the HAMMER in your bag; click on the red {Persian definition of Closet}

EXAMPLE 8 An excerpt showing an example of the input enhancement strategy using the Persian definition

Moreover, in the following example, Soheil, was the reader at that moment, and Artin enhanced the target word, *embers*, by pronouncing it loudly and spontaneously. After input enhancement, they searched the marginal gloss for its Persian meaning.

Artin	Soheil
Click on the **embers**	Click on the **embers**
امبرز علامت ؟	امبرز ؟ امبرز چی میشه ؟
نه بابا امبرز میشه ذغال داغ	علامت!
اینجا نتوشته روبروش	اون خو پلنکه

Artin	Soheil
Click on the **embers**	Click on the **embers**
Embers, sign?	Embers? What is **embers**?
No, that's {Persian definition}	Sign!
It's written there on margins	Isn't that for the *plank*?

EXAMPLE 9 An excerpt showing an example of input enhancement strategy by pronouncing the target word aloud

The exit interview confirmed the strategy. For example, when Minoo and Mähdieh were asked "What do you think you have learned about the words in this game? Meanings? Pictures? Spelling?" their answers emphasized the effect of the input enhancement strategy that they had used. They both mentioned pronunciation and definition as two salient aspects of the target words they had acquired after completing their digital game task.

What do you think you have learned about the words in this game? Meanings? Pictures? Spelling?	
Mähdieh	**Minoo**
بیشتر معنی و تلفظ	هم تلفظ ، هم معنیش
Definition and pronunciation more than anything else	*Both pronunciation and the definition*

EXAMPLE 10 The exit-interview data indicate the implementation of the input-enhancement strategy

In addition to the input enhancement strategy, the group A participants utilized a move called **planning**, as a problem-solving strategy. This move was also used by group C. In the move, after reading the game guide, analyzing the situation in the game and negotiating, the participants planned how to deal with and overcome a current problem. I labelled the move planning, as, unlike the group

B participants, neither the group A or C participants needed to guess the meaning of the target words from among the various options or need to create hypotheses about their definitions and test them. Thus, they could have planned their moves instead of hypothesizing them. For example, in the following example, Minoo and Mähdieh read the game guide and planned their further actions without any doubts or second thoughts.

Minoo	Mähdieh
Go out of the yard from the main door.	از حیاط برو بیرون
Click on the HAMMER in your bag	باید روی چکش کلیک کنیم
click on the door **sign**	علامته ، اینه علامت
Click on the nails and plank	Click on the nails and plank

Minoo	Mähdieh
Go out of the yard from the main door.	Go out of the yard
Click on the HAMMER in your bag	We must click on the hammer
click on the door **sign**	Sign, this sign
Click on the nails and plank	Click on the nails and plank

EXAMPLE 11 An excerpt showing an example of planning

5.2.2.2 Learning Approach by Group A Participants (Low Level of Involvement)

The analysis of the think-aloud data indicated a relationship between the universal moves and strategies employed exclusively by the participants in group A. In this section, I describe the strategy cycle used by group A in more detail and discuss what it reveals about the learning approach they applied in seeking to learn the target words.

The participants, that is, gamers, interacted, in the first place, with the text of the game guide in order to obtain the instructions. They then analyzed the digital game on the computer screen to find match the game guide instructions with the correct parts of the digital game. In other words, in their first move, they *searched for information*. In searching for information, the group A participants employed the *input enhancement* strategy to deal with the target words. Moreover, they *negotiated* over the information that they obtained from the game guide and their current situation in the digital game, in order to *plan* their future actions. They *took turns* to control the problem-solving process by either playing the digital game or searching for information in the game guide. After the success of their current plan in solving the problem, they *reviewed* their actions and completed tasks before moving to the next task. During the planning phase in the later stages of gameplay, they reviewed their previous actions and completed tasks to be able to reorganize their minds for finding the solution to their new problem. Their plans and actions were not always successful. In these cases, the participants performed their final option, i.e., *trial-and-error*. They clicked frequently on the objects in the scene in order to find the correct objects and solve the problem. If their trial-and-error moves failed, they either retried until they

succeeded or occasionally referred to the game guide. When successful, they reviewed the currently solved problem and the previous one to reorganize their minds and regain control of the digital game task.

Moreover, the internal elements of the digital game were a source of motivation for this group. In other words, aspects of the digital game such as story, sense of winning, and visuals motivated them to complete the tasks, progress in the game and solve more puzzles. This was confirmed by the participants when they were asked "What do you focus on when you play a video game? What about this game?" in their exit interview. This iterative pattern of the mental steps taken by group A for solving the challenges in the digital game task is illustrated in figure 18.

What do you focus on when you play video game? What about this game?	
Mähdieh	**Minoo**
روی اینه ببرم ، پیدا کنم چیزارو که به مرحله بعدی بریم ببینم آخرش چی میشه	برنده بشم ،این که کلمه ها رو که میگه بتونم پیدا کن میخواستم برم جلو ببینم چی میشه بعدش
To win, to find the objects and move to the next levels, to see the end of the story	*To win, to find the target words in the game, to find what will happen at the end*
Artin	**Soheil**
دنبال حل معما ،اون اسکلت کی بود توی اتاق	حل معما ، آخر داستان رو ببینم
Trying to solve the mystery, whose skeleton was that in the room?	*To solve the mystery, to see the end of the story*

EXAMPLE 12 Participants indirectly indicated the role of motivation in their exit interview

The reasons for the input enhancement strategy for dealing with the new target words employed by the participants in the group A can be speculated. One reason might be the type of instructions given. The task instructions, although not visible in the data, may have encouraged them to behave as they did during the task (Saravia, 1995, p. 30; Ericsson & Simon, 1987). The participants were told that the target words were presented in bold and that understanding them is important for proceeding in the digital game and completing the task. They might thus have developed a feeling for the bolded target words and prioritized them in their minds as very important in comparison with other new and unknown words in the game guide. Thus, they enhanced the bolded target words them by emphasizing, highlighting, and differentiating them from the rest of the words and the text either by pronouncing them loudly or by reading aloud the Persian definition in the English language guide. During the task, the group A participants never employed input enhancement for any other words than the target words in the game guide.

Another speculation is that the nature of the task as such prompted the group A participants to enhance the target words. Although this was not directly stated, the aim of the task was intentional vocabulary acquisition and the participants understood that they had to learn the target words by playing the digital game. Thus, their input enhancement strategy was an attempt at intentional learning. Employing the input enhancement strategy was an effort to

pay more attention to the target words. Furthermore, because the task was being performed cooperatively, the input enhancement strategy was either a hint or a signal to one's partner that the bolded words were important and needed more attention.

Yet another speculation concerns the pre-test effect. Although the participants were not informed that they were expected to learn the target words, exposure to the target words in the pre-test might have made the participants suspicious of them in the sense that they viewed the target words as critical. Hence, they deployed the input enhancement strategy to satisfy their feeling about the target words and to practice them by repeating and emphasizing them.

A final speculation could be the Hawthorne Effect, although I think this might not be as strong as the previous potential reasons for the use of input enhancement strategies by the participants in group A. The Hawthorne effect is defined as alteration in the behaviors of research subjects owing to their awareness of being observed (Seliger & Shohamy, 1989; Mackey & Gass, 2005). Thus, owing to the pre-test effect and their awareness of the recording devices, the participants alertness to the target words was heightened and they began to consider them as important to the researcher as well. Thus, they used the input enhancement strategy to indirectly notify the researcher that they were aware of the importance of the target words.

5.2.2.3 Group B (Moderate level of involvement load)

The group B participants showed the poorest performance of all three groups in the post-tests. Their task induced a moderate level of involvement load (index 9) and hence, theoretically, they were expected to show the second-best performance after group C. Surprisingly, however, this was not supported by the statistical results, which indicated group B's performance to be the poorest. Analysis of the group B participants' verbalized thoughts and online processes showed that, aside from universal moves, they employed an exclusive strategy and an exclusive move when they repeatedly encountered the target words. I labelled this strategy **inferencing from context**, and the move **hypothesizing**. In the strategy inferencing from context, the participants were actively *looking for contextual clues*. Here "context" refers to two contexts, i.e. the game guide text and the digital game. In applying the inferencing strategy in the game guide, the participants read the text and considered the surrounding textual information in order to guess at the best meaning. For example, in the following excerpt, Ali and Moein found the target word *closet* by searching for contextual clues.

Ali	Moein
Click on the HAMMER in your bag; click on the red **closet**	Click on the HAMMER in your bag; click on the red **closet**
رد یعنی قرمز ، اینو بزن برو برو برو یه لحظه بده {عوض کردن کنترل کننده}	

Ali	Moein
Click on the HAMMER in your bag; click on the red **closet**	Click on the HAMMER in your bag; click on the red **closet**
Red means {Persian definition} Go go go Give it to me {Mouse handed over}	

EXAMPLE 13 An example of participants employing inferencing strategies

The use of the inferencing strategy (textual clues) was also confirmed by the exit-interview when the gamers were asked "When playing a video game, do you use any tricks to help you with unknown words? What about this game?"

Do you use any tricks, when playing a video game, to help you with unknown words? What about this game?	
Ali	**Moein**
مثلا تو اولش که میخوندیم میگفت کلیک کن سر چیزی بعد مثلا میگفت رنگ زردیه ، قرمزیه ، کمک میکرد مثلا چیه ، تو تصویر نگاه میکردم چی قرمزه دنباله اونا میگشتم	مثلا شاید این کلمه تو جمله کاملی باشه ، از جمله هه دوتا چی دیگه بفهمم که وقتی که اون دوتا رو میفهمم پشتش اون چی کلمه جدیده هم میدونم در مورد چیه
For example, when we read, in the beginning, it said click on something; then, it said it was red, yellow; it was helpful to guess what it means; then we looked for red things only in the screen.	*Maybe, for example, a word is in a full sentence, if I know some words in that sentence, I can guess what the unknown word means.*
	Negar
	از متن واک ثرو استفاده می کردیم
	I used the walkthrough (the game guide) text.

EXAMPLE 14 Inferencing strategies mentioned by participants in exit interview

Seeking contextual clues also applied in the game context. Monitoring of the participants' actions in their in-game gameplay videos revealed that they hovered the mouse cursor over the objects on screen to test whether the object's name would appear. In this game, if the cursor is left on an object for almost 20 seconds, the game reveals the object's name. This feature of the game was also new to me. The exit-interview confirmed that guessing from the game context by hovering with the mouse cursor showed that the inferencing from the context strategy was a successful effective for the group B participants.

Do you use any tricks, when playing a video game, to help you with unknown words? What about this game?	
Tara	**Negar**
روی اجسام روی بازی می رفتیم ، اسمشون رو می نوشت ، می فهمدیم منظورش چیه.	روی شکل ها می ایستادیم تا کلمه ها ظاهر بشن.
We hovered the mouse curser over the game objects, it showed the name, then we could understand what it means.	*We hovered the mouse over the objects to show the name of the objects*

EXAMPLE 15 Inferencing strategies used by participants in the game context

The Group B participants also performed an exclusive move, i.e. hypothesizing. This move was labeled hypothesizing as, for each target word, the participants selected one of three options. Thus, they hypothesized their future actions in every encounter with the target words. They manipulated and restructured their hypotheses using contextual clues and the inferencing strategy. For example, for the word *sign*, Ali and Moein had three options. Thus, based on those three options they developed hypotheses that determined their future actions.

Ali	Moein
	Click on the HAMMER in your bag, click on the door **sign**; Click on the nails and plank
	میگه روی چکش کلیک کن ، بعد کلیک کن روی در...
ذغال داغ ، علامت ، حصار فلزی رو حصار فلزی بزن ببنم	
	{کلیک ... کلیک}
خوب الان ما فهمیدیم حصار فلزی نیس ذغال داغ... بزن بزن ، اینو بزن	

Ali	Moein
	Click on the HAMMER in your bag, click on the door **sign**; Click on the nails and plank
	It says click on the hammer, then on the door …
Ember, sign, or fence Click on fence to check	
	{Click … click}
Now, we know that it is not fence Check with embers Click, click on this one.	

EXAMPLE 16 Participants' use of the hypothesizing move by participant

Thus, through the hypothesizing move and testing their hypotheses, they eliminated the options for each target word until found the correct one.

5.2.2.4 Learning Approach by Group B Participants (Moderate Level of Involvement)

As with group A, analysis of the data showed that the group B participants' moves and strategies were dynamically related and indicated a learning cycle. In this section, I describe group B's cyclical employment of exclusive strategies and

universal moves and discuss how this cycle formed the group B participants' learning approach.

The participants in group B first *searched* for in the game guide in an effort to *infer* the correct meaning of the target words. They also searched for contextual clues in the game by hovering their mouse cursors on the objects in the scene. After their searches, they *negotiated* on their guesses and developed *hypotheses* for their future actions. They then tested their hypotheses in *turns* although turn-taking was not done in a pre-arranged order. Sometimes, turn taking did not happen at all. If their hypotheses were successful, they *reviewed* their current and previous actions to organize their minds and proceed to the next task. In the case of failure, they either tested the rest of their hypotheses, referred to the contexts of the game guide and the game to develop new hypotheses or they relied on the *trial-and-error* move as their final option. As with group A, the digital game was also a motivator for the group B. this was found when the group B participants were asked "What do you focus on when you play video game? What about this game?" The iterative pattern in the mental steps taken by the group B participants to solve the challenges in the digital game task is illustrated in figure 19.

What do you focus on when you play video game? What about this game?	
Ali	**Moein**
پیروزی	بیشتر دور و اطراف که اگر یه چیز مشکوکی بود به چشمم بخوره استفاده کنم تا رد کنیم مرحله رو
To win	*I was looking carefully into the scene to find suspicious things and use them to solve the problem*
Tara	**Negar**
معماهای بازی رو حل کنیم و بریم جلو	حل معماها و باز کردن درها و ببینیم چی میشه
To solve puzzles and progress in the game	*Solving the problems and puzzles, opening the doors, to see the ending*

EXAMPLE 17 The motivational role of the digital game indicated by participants in their exit-interview

Reasons can be proposed for the group B participants' employment of inferencing strategies and the hypothesizing move, which formed their learning approach. The first reason might be lack of access to external sources such as dictionaries, internet etc. They were engaged in a task with unknown words and three marginal definitions for each word. Given their lack of access to facilities like dictionaries, internet etc., the participants had no choice but to makes inferences and guesses, while testing their hypotheses was an inevitable next move in finding out if their guesses were correct (Hulstijn, 2001; Rott, 2005).

The second reason might be the nature of the task. As already described, the group B participants had to infer the best definition from the context of use in the game. Therefore, unintentionally, they were encouraged to make inferences and to hypothesize.

The third, and final reason might be the instruction given. The participants in group B were asked to read the game-guide text, select the correct definition, solve the problem, and proceed in the game. To do this, they might have

employed the inferencing strategy and hypothesizing move to follow the instructions as they understood them.

5.2.2.5 Group C (High level of involvement load)

Group C experienced learning the target words by interacting with the digital game task, a procedure which induced the highest level of involvement load (Index 12). Based on the statistical findings, group C had the highest mean rank in the post – tests and outperformed groups A and B in both post-tests. Scrutiny of their approach to the digital game task yielded two main recurring themes that could be categorized as exclusive strategies: these were labeled **memory search** and **feedback request**. Moreover, like group A, the group C participants also implemented the **plan** move.

The participants in group C attempted **memory search** when they encountered the target words for the first time. For example, when they met the word *latch* for the first time, both participants kept silent and started thinking. They thought aloud as follows:

Shirin	Shadi
click on the latch	
لَچ چی بود؟ {سکوت} ها! یعنی چفت در	لَچ چی بود؟ لچ ، لچ ، لچ چفت در!

Shirin	Shadi
click on the latch	
What was the latch? {Silently Thinking} Aha! It means (Persian Definition)	What was the latch? Latch, latch, latch (Persian Definition)

EXAMPLE 18 Participants employing the memory search strategy

This was one of their very first encounters with one of the target words. The same process occurred for most of the other target words. This excerpt shows both participants engaged in an in-depth memory search. Their memory searches displayed two frequent patterns: thinking silently and gazing at either the screen or the game guide and repeating the word that they were thinking about.

However, in some situations the participants either were not sure enough of the outcome of their memory search or were incapable of recalling the Persian definitions. In this case, they relied on their partner's knowledge and **requested feedback.** For instance, when trying to find the word *hook* in the game, Shirin was not sure of what she had recalled about the word. She requested feedback from her partner Shadi to either confirm or reject the recalled Persian definition.

Shirin	Shadi
هوک چی بود؟ {سکوت} مطمعن نیستم ، یعنی قلاب؟	هوک؟ آره فک کنم میشد قلّاب

Shirin	Shadi
What was the hook? {Silently Thinking} I am not sure, does it mean (Persian Definition)?	Hook? Yeah, I think it was (Persian Definition)

EXAMPLE 19 A participant's request for feedback (confirmation)

In the case of the word *Portrait*, for example, Alireza was unable to recall anything. He asked Ali for the definition.

Ali2	Alireza
خوب گفت Go back down.	Click the HAMMER in your bag, click on the portrait
عکس	پورتریت چی بود؟

Ali2	Alireza
Well, it said Go back down.	Click the HAMMER in your bag, click on the portrait
{Persian Definition}	What was the portrait?

EXAMPLE 20 A participant requesting feedback (recall)

The exit-interview data validated these findings. When the participants in group C were asked "What did you do if you couldn't remember/recognize the meaning of the unknown words in the game guide?" they answered by trying to recall them or by asking their partners for help. In below, you can find their answers:

What did you do when you met an unknown word/selected words in the game guide of this video game?	
Shirin	**Shadi**
یا مثلا مجبور بودم خیلی فک کنم تا یادم بیاد یا ا شادی میپرسیدم که زودتر پیش بریم	مثلا ا همشون یه کوچولو تو ذهنم بود مثلا دیگه وقتی متنشو میخوندم میدیدم اه این کلمه گفته بیشتر با توجه به بازی بعضی وقتا هم شیرین
I must either think a lot about that to recall its definition or ask Shadi to progress	*For example, I have a small piece of information about any of them. Therefore, when I read the instructions and I saw the words, I remembered them. Otherwise, I asked my partner, Shirin.*
Ali2	
از پارتنر کمک می گرفتیم	
We asked partner for help	

EXAMPLE 21 Memory search strategies mentioned by participants in their exit-interview

The group C participants, like those in group A, planned their future actions. This move is considered planning, rather than hypothesizing, as the participants had been exposed the definitions at least once before they began their main task. Thus, they might have planned their actions rather than hypothesizing. For instance, in the following excerpt, Shirin and Shadi are planning their moves.

Shirin	Shadi
Take the STONE and RING بعدش	
Click on the old shack	
خوب ، کلبه اینه ، آره	
بعدش Use the STONE and click on the window 2 times	
{کلیک... کلیک}	خوب

Shirin	Shadi
Take the STONE and RING Then,	
Click on the old shack	
Ok, {Persian definition} This one, yes	
Then, Use the STONE and click on the window 2 times	
{Click … Click}	Ok, good

EXAMPLE 22 Participants planning their future actions

Moreover, regarding unknown words, other than the target words bolded in the game guide the group C participants used a strategy, which could be labeled the **word association** strategy. That is, they referred to the semantic association of contextually known words to infer either what an unknown word meant or which on-screen object they had to use. The following excerpt from exit-interview exemplifies this strategy.

What did you do when you met an unknown word/selected words in the game guide of this video game?	
Ali2	**Alireza**
بغیر از اینایی که خونده بودیم ، بقیه رو با آزمون و خطا ، و خوندن بقیه متن و در نظر گرفتن کلمات و اشیا دیگه پیدا می کردیم.	مثلا وقتی بیلچه رو داشتیم و نمی دونستیم چی کارش کنیم ، جایی که به خاک میرسیدیم ، حدس میزدیم «آهان ، باید از بیلچه استفاده کنیم»
Other than the words we read, we found the new words by trial-and-error, reading the rest, and considering the other words and objects.	*For example, when we didn't know what to do with the shovel, when we found soil, we guessed "yeah, we must use the shovel".*

EXAMPLE 23 The word association strategy mentioned by participants in the exit-interview

5.2.2.6 Learning Approach by Group C Participants (High Level of Involvement)

For group C, exploring the data as a whole and analyzing them from a wider perspective indicated a pattern that generated a distinctive learning approach. Moreover, unlike in the other groups, the holistic analysis of data showed a dichotomous relation between memory search and a feedback request. In other words, when memory search was conducted feedback requests were absent. In contrast, feedback requests were prevalent when the memory search had been unsuccessful. Therefore, with respect to universal moves and exclusive strategies, the mental steps that participants followed in performing the digital game task, showed a pattern that can be described as follows. In this group, as in group A, the main source of *information search* was the game guide. When the participants encountered the target words, they either employed *memory search,* in the first place, or *requested feedback.* If there were other unknown words in the text, they tried to infer them using the *word-association* strategy. After interacting

with the game guide and analysing the situation in the game, the participants *negotiated* in order to develop *plans* for their future actions. If, after implementing their plans, they succeeded, they *reviewed* their previous actions and moved to the next task. In the case of failure, they *took turns* to read the game guide and play the game for implementing their upcoming actions again. They rarely tried solving problems or finding the bolded target words by *trial-and-error*; however, as mentioned, they occasionally employed this move for other unknown words. The digital game was also a source of motivation for the participants in this group. For instance, when asked what they had focused on the most during their interaction with the digital game their answers indicated the motivational appeal of the game. Thus, the manner in which the digital game presented problems encouraged the participants to solve them can be viewed as the motivational factor that powered the iterative cyclical pattern for solving each task. In other words, the digital game made problem-solving interesting by via internal elements, such as the story.

What do you focus on when you play video game? What about this game?	
Shirin	**Shadi**
میخواستم ببینم آخرش چیه میشه	توی این بازی هم بیشتر سعی کردم چیزایی که میخواد رو پیدا کنم تا ببینم چی میشه
I wanted to see the ending and find out what will happen	*I was trying to find the clues that the game needed us to find to see the end of the story.*
Ali2	**Alireza**
ته بازی چی میشه	اینکه تهش چی میشه ،پشت اون در قرمز گاراژ چی بود
What happens in the end	*What happens in the end, what was behind the red garage door?*

EXAMPLE 24 The digital game as a motivational source indirectly referred to in the exit interview

This iterative pattern found as the mental steps taken for solving the challenges in the digital game task by group C is illustrated in figure 20.

There are many possible reasons for the use of memory recall, feedback request, word association strategies, and planning by group C. For instance, in the first place, as in the other groups, the group C participants had no access to external sources such as dictionaries or the internet. Hence, their reliance on their memories and recalling the definitions was inevitable. They had no choice but to dig into their memories to solve the target word-related problems in the game.

Another reason might be the nature of the instructions. The participants, in this task, were asked to memorize the definitions of the target words. Moreover, they were told that they would need to recall them during gameplay in order to proceed in the game. To do this, they might have prioritized memory search and memory recall as their first strategy to overcome the challenge of knowing a target word.

A further possible reason is the cooperative nature of task performance. As already mentioned, the lack of external information sources confined the participants to searching their own memories. Moreover, each partner was aware that the other had tried to memorize the word list and the definitions. This idea may have persuaded them to trust each other's memory and request feedback

from each other. This is supported by the fact that they called the researcher for assistance very few times throughout the gameplay session. Moreover, because the task was done cooperatively, one's partner was also considered a source of information to be consulted, when the memory recall strategy had failed. Thus, lack of access to other sources such as the internet and dictionaries, the nature of the instructions, and cooperative task accomplishment were factors motivating them to compensate for their memory failures by requesting feedback.

A final speculation concerns the style and structure of the game guide. Unlike the other groups' game guides, group C's game guide contained no definitions, glosses, in the margins. Therefore, in addition to the influence of the instructions given, the game guide itself was a persuasive factor in encouraging the participants to search their memories, request feedback and use associative word strategies.

It is noteworthy that group C performed the actions and turn-takings in an orderly manner. In other words, because their problem-solving strategies, such as turn-taking, negotiating, and planning were well organized and not overwhelmed by a flood of plans and hypotheses, they had a better understanding of the text. Thus, trial-and-error was a rare move in this group.

5.2.2.7 Learning Approach Models

The dynamic relations between moves and strategies are illustrated in the following flowcharts. It should be noted that the learning approaches were iterative in every stage of problem solving in the digital game task In the figure 20, the trial-and-error move is shown by dashes, for the group C, in order to indicate that it was rarely applied for the purpose of finding target words.

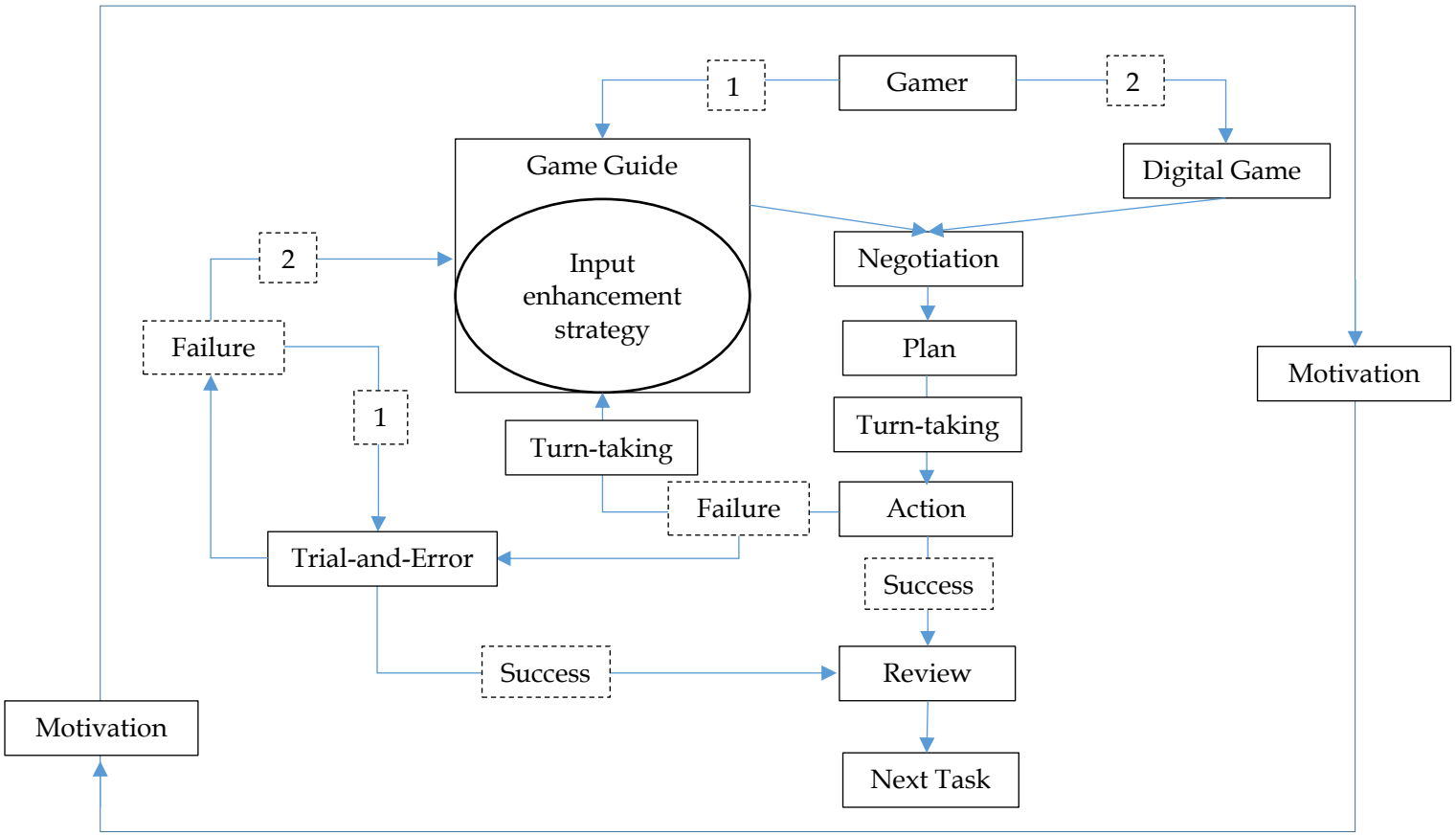


FIGURE 18 The Group A Learning Approach

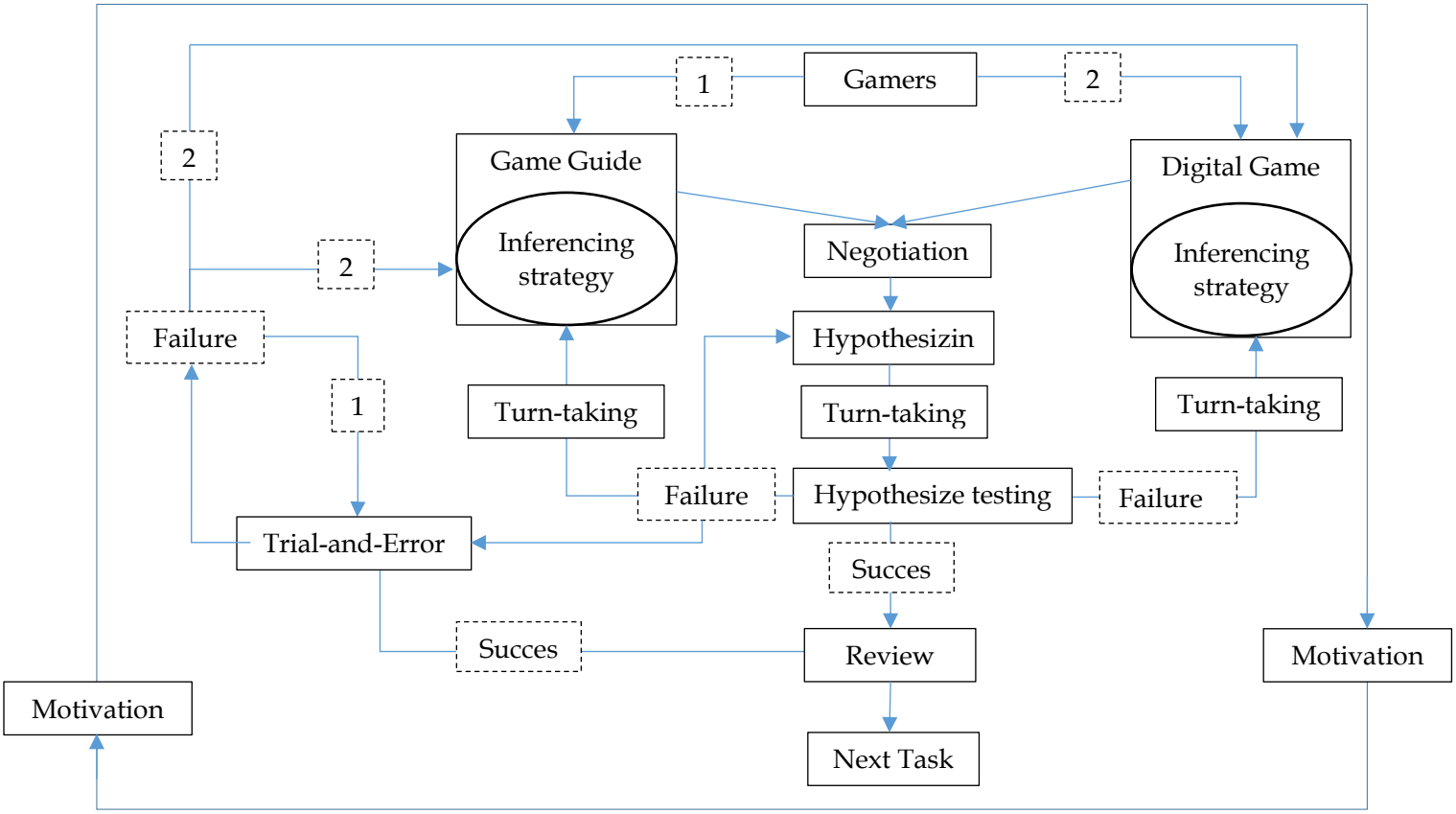


FIGURE 19 The Group B Learning Approach

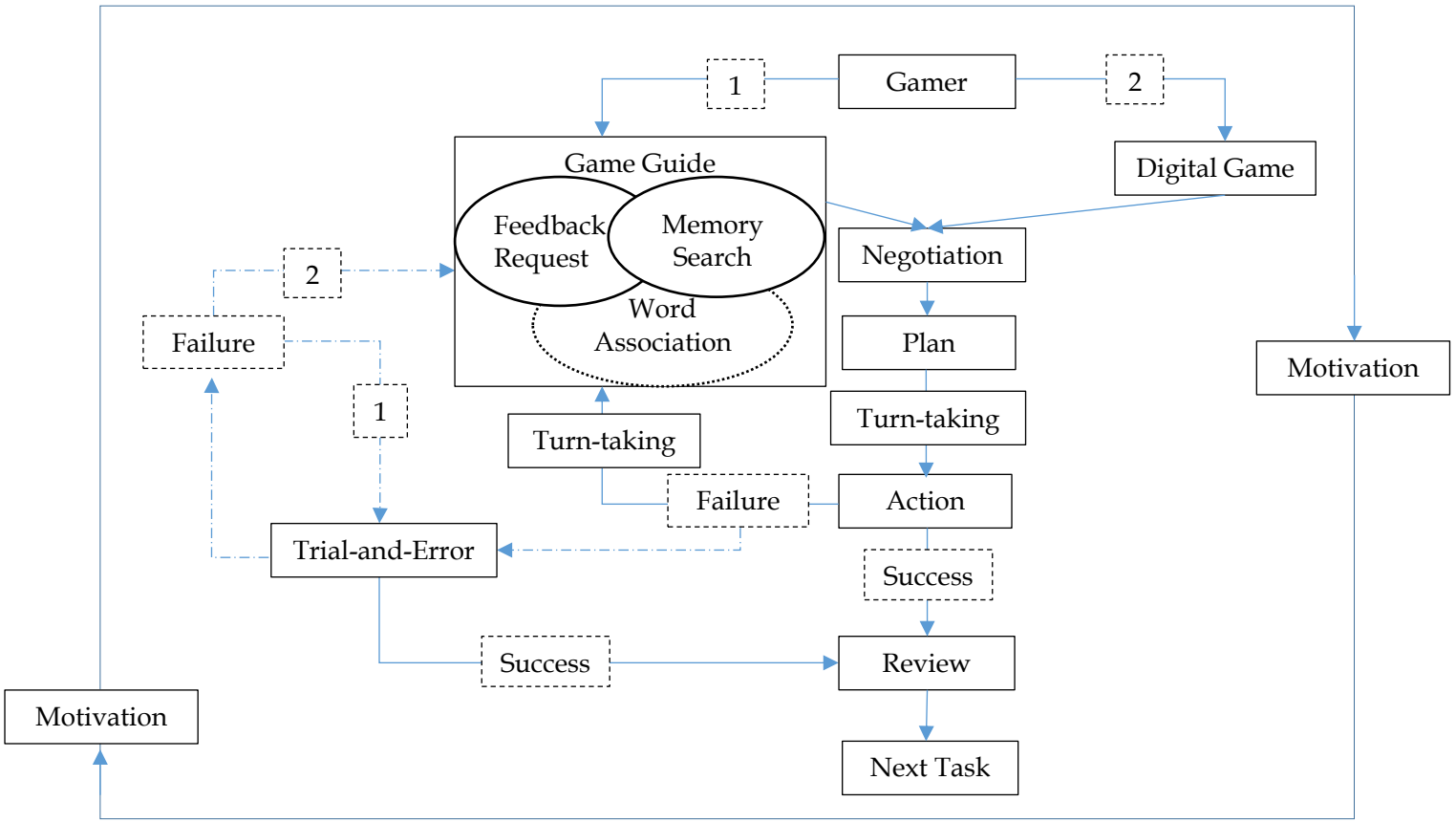


FIGURE 20 The Group C Learning Approach

5.2.3 Comparing and Contrasting the Emerged Categories (Moves and Strategies)

Content analysis of data has indicated interesting factors that have played significant roles in groups' performance for the digital game task accomplishment. Besides, these factors have the potential to reveal important information for the research if they are compared and contrasted. Because "Comparing has the capacity to reveal the link between codes […] it should rather capture something important in relation to overall research" (Vaismoradi, Jones, Turunen, & Snelgrove, 2016, p. 105). Thus, comparing and contrasting those categories might elevate our understanding of the outperformance of the group C, the poor performance of the group B, and why the quantitative results did not support the theory. The performance of the groups can be compared on their universal moves and exclusive strategies. Therefore, the comparison was done in two dimensions and from two different perspectives.

In the first dimension, the groups showed noticeable differences in the role of the universal moves. One way to compare themes and categories is by their frequency of occurrence in the coded data (Vaismoradi et al., 2016). Thus, in this dimension, I compared and contrasted the groups by the frequencies of the different universal moves.

TABLE 22 Table of Frequencies for the Categories of Universal Moves

	Group A		Group B		Group C	
Information Search	156	42%	106	27%	132	42%
Negotiation	81	22%	150	39%	59	19%
Turn-taking	38	10%	61	16%	43	13%
Trial-and-Error	40	11%	49	13%	24	7%
Review	56	15%	18	5%	59	19%
Total	**371**		**384**		**317**	

The frequency table shows a significant between-groups difference in the *information search* move. Groups A (156) and C (132) searched for information more frequently than Group B (106). In other words, A and C showed a more intensive engagement with the context, allowing them more exposure to the text and the target words. Although *information search* was performed in both the game guide and game, the recorded video clips showed that the game-guide text was primary for seeking information, especially for groups A and C. Thus, because the group A and C participants experienced more engagement with the text, their chance of exposure to the target words was higher, and then, which also raised their possibilities of acquiring the target words. This inference is in line with previous findings emphasizing the role of frequency of exposure and its significant effect on the acquisition of new vocabulary items (Huckin & Coady, 1999; Nation, 2001; Webb, 2007b). The frequencies of the *review* move also support the above reasoning. Groups A (56) and C (59) reviewed their previous actions significantly more often than the group B (18). Hence, the frequency of exposures to the text and target words was higher in groups A and C. Piirainen-Marsh and Taino (2009) suggest that these types of exposures encourage language learners to analyze the second language more deeply, thereby revising and developing their linguistic competences. Thus, the frequency of exposures to the target words might have been a defining and significant factor in superior performance of groups A and C in comparison to that of group B.

Furthermore, group A had probably outperformed group B because their task accomplishment approach was almost the same as that of the group C. The superiority of group A performance in task accomplishment and the post-tests might also be due to their more frequent exposure to the target words. However, in comparison to group C, group A's performance was poorer. This difference may also be explained by exposure frequency. Owing to group C's task design, this group was exposed to the target words before they addressed their main task. Thus, it possible that group C's frequency of exposure to the target words was greater than that of group A. Moreover, they had studied the target words two days before commencing their main task, a procedure known as the spaced-repetition effect. In the literature, spaced-repetition has been shown to be a significant factor in acquiring new vocabulary items; furthermore, the effect is even superior to that of high frequency of exposure (Nation, 2001; Nation & Webb, 2010). According to the Technique Feature Analysis checklist (Nation & Webb, 2010), spaced-repetition is one of the factors that increases the task-

induced level of involvement load. Thus, due to spaced-repetition and high frequency of exposure, the group C participants were more involved with the target words than the other two groups. It can be assumed, therefore, that their memory traces of the target words were deeper and consequently that their target word acquisition was superior. Thus, the *information search* and *review* moves might have been the contributory factors that leveled the differences in performance between the groups in the digital game tasks.

Another move which can be considered as a factor contributing to differences in performance between the groups is *negotiation*. According to the table, *negotiation* was iterated far more often by group B (150) than groups A (81) and C (59). This higher frequency might, for two main reasons, also explain the poor performance of the group B. First, the higher number of negotiations might have produced too many options owing to the participants' uncertainty over their choices of definitions (Rott, 2005). Thus, it can be argued that the participants did not have a clear understanding of the game-guide text, as also found by Rott (2005) in the case of multiple-choice gloss readers. Thus, in their efforts to understand the text and proceed further in the game, the participants worked with too many ideas and hypotheses. By the same token, since the text and its comprehension were their major problems, they might not have had enough time to consider and to process the target words. They may also have carefully searched the game guide and the game for information and to control and balance the flood of ideas, thereby hampering the development of new ones before they had tested their currently developed online hypotheses. Consequently, they relied on negotiating and then hypothesizing more than on the other universal moves. A reason that suggests this interpretation is the frequency of the *trial-and-error* (49) move, which was less than expected given the participants' uncertainties about the definitions. Therefore, owing to their mental load, lack of proper comprehension of the game guide text, and the large amount of online ideas and hypotheses, they did not employ the *trial-and-error* move as often as expected.

The second potential main reason for the group B's lower number of exposures to the text is that they relied more on their negotiations and follow-up hypotheses for solving the problems in the task than on the information search move and the game guide, which were discussed as the probable factors r boosting the performance of group C. Therefore, the negotiating and hypothesizing possibly negatively affected group B's vocabulary acquisition negatively. Several reasons can be advanced for this. First, because the group B participants engaged deeply in negotiation and hypothesis development, they might have skipped or overlooked the form of the target words in the game guide text (Rott, Williams, & Cameron, 2002). In other words, because all the three definition options were there in the margin then, why would they bother with the form of the target word? Had they done so, they might, after negotiation, have developed three different hypotheses and solved the problem by just testing them. Thus, the word form–meaning link might not have even formed; or maybe it formed loosely, which might explain their poor performance in the post-tests.

In contrast, for the group C, the development of the word form-meaning link was essential because they needed to recall the definitions based on the target word forms. Consequently, they might have invested a lot of attention in the target forms and might have endeavored to strengthen the link so that they could recall the definitions and successfully complete the task. Furthermore, like group B, group A might have also overlooked the target word forms; however, their performance was better because their minds were probably not so busy storing and processing numerous ideas and hypotheses during their engagement with the game guide text. It is possible that just one exposure to the target word forms was enough for them to develop the form-meaning link. Their performance was poorer than that of the group C, also because they had probably developed a weaker form-meaning link than group C.

In addition to word-form skipping, the nature of the task might be another reason for the performance differences in this study. The nature of the task could lead to the unnecessary depletion of attentional sources (Rott, Williams, Cameron, 2002). In other words, the higher frequency of negotiation moves in group B might indicate that the participants probably invested very little attention in learning the target words because the nature of the task encouraged them to make inferences, negotiate and hypothesize. As a result, they might have depleted most of their attentional sources, thereby impairing their drawing of successful inferences the usefulness of their negotiations, and the effectiveness of their hypothesizing. As attention and observation are the preliminary stages of learning (Truscott, 1998), the group B participants probably did not effectively learn the target words because did not invest enough attention in the target words. In contrast, the nature of the group C's task might have persuaded the participants to invest most of their effort and attention in learning the target words, leading them to a better learning outcome than that of group B.

The second dimension of comparing the groups was addressed the groups' different vocabulary learning strategies yielded by the qualitative analysis. It is possible that the specific strategies the participants had employed to deal with the challenges presented by the target words explained their relative success or failure in acquiring the target words.

In this study, the three groups applied different strategies, for solving the target word problem in their digital game tasks. In the same vein, Rott (2005) compared vocabulary acquisition tasks and investigated the effect of first-language single glosses and first-language multiple-choice glosses on vocabulary acquisition. She also identified the participants' strategies when they approached the text in both conditions using a concurrent think-aloud protocol. She also considered level of involvement load and found that the multiple-choice group was far superior in immediate vocabulary gain and in retention four weeks later despite their poor performance in comprehension of the text. She discussed her findings with respect to strategy use and level of involvement load. She observed that the multiple-choice group employed strategies like inferencing and hypothesis-testing while the single gloss group simply used a meta-cognitive strategy, i.e. glancing at the definitions in the margins. She suggested that the

multiple-choice gloss group experienced a higher level of involvement load in their vocabulary acquisition task. Moreover, the *search* and *evaluation* components of the level of involvement load hypothesis indexed as greater; consequently, the traces of the new target words were deeper, the form-meaning link developed more strongly, and hence the group B participants learned the target words better. In contrast, the group A participants only received scores for the *need* component of the level of involvement load hypothesis; thus, the form-meaning link was markedly weaker for group A with only single glosses, and hence their vocabulary acquisition was inferior. These findings had previously been reported by Fraser (1999), who found that the inferencing strategy was a superior and efficient lexical processing strategy for learning unknown words in reading contexts. However, in the present study, the results indicated a reverse, as group A was superior to group B. The contradiction between the present and Rott's (2005) findings may be explained by strategy selection, as discussed by Gu and Johnson (1996). According to Gu and Johnson (1996), the type of strategy that language learners choose for learning vocabulary items depends highly on the person, context and task at hand. Therefore, the main reason for the discrepancy between the results might be the context. Although the present task implicitly prompted the participants to employ the same strategies, owing to the multimedia-enhanced context rather than text-only context, the rate of vocabulary acquisition was poorer in group B despite the fact their route was the same as that of Rott's (2005) participants.

Why the multimedia context reversed the outcome merits discussion. According to deHaan, Reed, and Kuwada (2010), the process of vocabulary recall is hampered by cognitive overload. In their study, the participants, whose tasks demanded a higher level of interactivity, experienced cognitive overload. For this reason, they recalled fewer words than their counterparts in the low-level-interactivity-task group. Based on this finding, the inference may be drawn that the digital game task and context might also have overloaded the participants' minds in the present study.

> The objective of [cognitive Load Theory] CLT is to predict learning outcomes by taking into consideration the capabilities and limitations of the human cognitive architecture […] taking into account the demands on cognitive resources induced by the complexity of the information to be learned, the way in which the instruction is presented to the learner, and the learner's prior experience and knowledge, CLT aims to predict what makes learning successful and how learning can be effectively supported by teaching and instruction. (Plass, Moreno, & Brünken, 2010, p. 1)

In the same vein, Mayer and Monero (2003), after twelve years of experimentation, emphasized the critical and defining role of CLT in learning from multimedia materials. Moreover, they discussed nine distinctive sources that may cause cognitive overload and thus hinder learning in multimedia learning contexts.

Thus, drawing on Mayer & Moreno (2003), it is worth discussing group B's poor performance can be reviewed and discussed, from the perspective of the sources of cognitive overload. Group B's task and choices of strategies may have

had features that could cause cognitive overload, such as the *split-attention* effect (Mayer & Moreno, 2003). The group B participants had to read the game-guide text, understand the instructions, and apply them in the digital game. In other words, they had to split their attention between the game guide and playing the digital game. However, the presence of multiple-choice glosses possibly caused them to further split their attention between these and other sources of information, such as their partners' ideas and the digital game events, while simultaneously running processes such as inferencing, negotiating, hypothesizing, and hypothesis-testing. Consequently, because the processing of multiple elements may have induced the split-attention effect and overloaded their working memory overloaded. Thus, first, this would hamper their understanding of the game guide; second, it would hinder their task performance; third, the new information, that is, the target words, would not be sufficiently processed in their working memory; and finally, it would slow down the process of acquiring the target words.

Although the split-attention effect was probably inevitable and potentially hindered acquisition of the target words for groups A and C, it was probably prevented by their *signaling* (group A) and *pre-training* (group C) techniques,(Mayer & Moreno, 2003). In the signaling technique, cues must be provided to guide participants how to process the material effectively (Mayer & Moreno, 2003, p. 46). Group A's marginal glosses possibly acted as such cues and elevated their understanding of the game guide text. Therefore, they did not, like group B, need to split their attention between these and other information sources and invoke processes for guessing, hypothesizing, and hypothesis-testing. Thus, their cognitive processes might have been beneficial and effective for the development of the form-meaning connection and subsequent acquisition of the target words in the digital game-based vocabulary acquisition context. For the group C, the pre-training technique, as the name indicates, probably assisted the participants in how to process the new information, i.e. target words in the task. Thus, the group C participants, who had processed the target words at least once before they embarked on their main task, might not have needed to invoke extra processes for finding the target words in the game guide because this was no longer a challenge for them; therefore, they did not need to split their attention between information sources other than the game guide and the digital game. Moreover, they might also have used the target words and pretraining to better focus their attention on memory recall and feedback request strategies, which were more important for accelerating retention of the target words. Thus, groups A and C might have been more successful in the post-tests because their tasks and choices of strategies possibly led them to implement the appropriate procedures, within the capacity of their working memory, for target word retention in a multimedia learning context without the distraction of attending to unnecessary sources of information. In contrast, group B's task and choices of strategies might have invoked improper or additional processes that overloaded the participants' working memory by splitting their attention between

unnecessary information sources, and hindering their learning of the target words.

Finally, based on the discussion by Gu and Johnson (1996) on the selection of vocabulary learning strategies , the task itself was another effective factor in group B's choice and deployment of these strategies. Group B's task encouraged inferencing, guessing and hypothesis-testing (Hulstijn, 2001), probably because of the multiple-choice glosses and the game guide instructions. Moreover, they had the opportunity to employ them in two interrelated contexts, i.e., the game-guide text and the digital game. Thus, the guessing strategies that were employed in the digital game context, like *hovering the mouse* in addition to features of the digital game such as immediate feedback (Kiili, 2005), gave them, unlike the other two groups, an opportunity to avoid deeper analysis of and interaction with the target words and the text. Hence, the participants might either have skipped the target forms and just hypothesize-tested the multiple-choice options in the margins of their game guide by abusing the immediate feedback feature, hovered their mouse cursor over objects at random in the hope of hitting on the right one. By abusing those strategies, they might not even have made the form-meaning connection and not started the vocabulary acquisition processes at all. Thus, another speculation is that malperformance of the task and its diversion from its intended design and purpose – despite its adequate involvement load – and its tendency to encourage the use of improper strategies were probable factors in the poor performance of the group B participants in the post-tests.

In Rott's (2005) study, both the choice and quantity of strategies used were the defining factors in effective vocabulary gain in vocabulary acquisition tasks. Accordingly, it can be inferred that, compared to the other two groups. the group C participants might have employed strategies that improved the rate and quality of vocabulary acquisition in the digital game-based vocabulary acquisition task. Moreover, group C's strategies might also have limited cognitive overload by trimming processes to suit the participants' working memory capacities. The group A participants' strategy choices had nearly the same attributes as those of the group C participants, yet their strategies were not as effective. Thus, choice of strategies and task design, which might affect the strategy choice may explain the superior performance of group C and the poor performance of group B.

6 DISCUSSION AND CONCLUSION

It has been argued in the vocabulary acquisition literature that level of involvement load hypothesis (ILH) is a fairly reliable index that predicts the rate of vocabulary acquisition tasks (Hulstijn & Laufer, 2001). This notion has been put into practice by researchers and, despite mixed results on the predictive success of the ILH (Keating, 2008; Yaqubi, Rayati, & Gorgi, 2010; Kim, 2011; Tahmasbi & Farvardin, 2017; Zou, 2017), it is considered an important factor in vocabulary acquisition tasks (Huang, Willson, & Eslami, 2012). In turn, digital games, a relatively new phenomenon, are also considered a helpful learning aid (Kirriemuir, 2002).It is, therefore, unsurprising that language teachers have longed integrate digital games in language learning tasks in order to boost learning of such aspects as vocabulary. However, digital games supply us with new and uncharted contexts, features, and tools that can either hinder or accelerate vocabulary acquisition. The role of task features, especially the effect of involvement load on digital game-based vocabulary acquisition tasks is largely uncharted territory. It is, therefore, necessary to explore and studied digital games before using them in vocabulary acquisition tasks. The present study investigated the effectiveness of different task-induced levels of involvement load on the acquisition of target vocabulary items in a digital game-based vocabulary acquisition context. Overall, the outcomes of this study, as stated in the previous chapter, indicated that digital games are effective on the acquisition of different dimensions and scopes of target vocabulary items, such as form, meaning, reception, production, recognition and recall. Furthermore, such games better accelerate the productive than receptive dimension of vocabulary acquisition. However, the findings of this study also partially support the premise of the ILH. In other words, the results showed that although the digital game task inducing the highest level of involvement load enhanced the acquisition of the target vocabulary items both qualitatively and quantitatively, the outcomes of the other tasks were not in the predicted order. The task with moderate involvement load did not assist the participants to acquire the target words while the task with the lowest involvement load did. Moreover, the qualitative analysis of the effect indicate that the effectiveness of the ILH, in

digital game-based vocabulary acquisition tasks, is conditional and depends on factors such as the manner of vocabulary presentation, the strategies used for completing the tasks, and the digital game context and tasks. In line with some of the previous studies (Folse, 2006; Martínez-Fernández, 2008; Yaqubi, Rayati, & Gorgi, 2010; Jahangiri & Alipour, 2014; Zou, 2017), the findings of this study challenge the predictive precision of the ILH and emphasize the complexity of vocabulary acquisition.

In this chapter, the findings of this study are discussed and explained in relation to previous studies and findings in the vocabulary acquisition literature. Furthermore, possible pedagogical implications and directions for further research are suggested. To fully discuss the findings of this study, I have divided this chapter according to the research questions. In this study, in addition to the findings on the ILH, other observations were made about the effect of digital games on the acquisition of target vocabulary items. In the first section, I discuss the effect of the digital game on the acquisition of the different aspects, dimensions, and scopes of the target words. In the second section, I discuss which dimensions and scopes of the target words were effectively acquired through interaction with the digital game tasks. Finally, I address in detail the major object of this study, i.e., the effect of different levels of involvement load on the acquisition of target vocabulary items in a digital game task.

6.1 Research Question 1: What is the effect of the digital game, in different levels of involvement load, on the acquisition of target vocabulary items?

First, I aimed to investigate the effect of the digital game on the acquisition of target words; that is, to find out whether the novel mental construct, or in this instance new vocabulary item, is the outcome of participants' interaction with the digital game tasks. I consider this a must-be-done analysis in every digital game and vocabulary acquisition study owing to the multidimensional nature of word knowledge: "the effect of DGBL on vocabulary learning may vary with game design feature" (Chen, Tseng, & Hsiao, 2018, p. 73). The findings indicate that the present digital game tasks promoted vocabulary acquisition, as the participants showed improved performance in their delayed post-tests. Moreover, the results support previous findings (Rankin, Gold, & Gooch, 2006a; Yip & Kwan, 2006; Muhanna, 2012; Chiu, Kao, & Reynolds, 2012; Vahdat & Rasti Behbahani, 2013; Chian-Wen, 2014; Ebrahimzadeh & Alavi, 2017; Chen, Tseng, & Hsiao, 2018). This study adds to the literature in demonstrating the effectiveness of the digital game tasks on the acquisition of such dimensions and scopes of word knowledge as production, reception, recognition, and recall, as well as form-meaning connection. This outcome can also be considered an important contribution to the vocabulary acquisition literature because the nature of word knowledge is multidimensional, and hence the acquisition of either form or meaning alone

cannot be considered adequate vocabulary acquisition. In other words, based on this new finding, it seems that digital games have the potential to overcome the difficulty of acquiring word knowledge, which is both an interesting and debatable notion.

The effectiveness of digital games on the acquisition of aspects of word knowledge can be discussed from both cognitive and motivational perspectives. From the cognitive perspective, the role of context in the process of encoding new lexical items is paramount. In the vocabulary acquisition literature, especially in the incidental model (Hulstijn, 2001), context is one of the major contributors to the degree of success in the encoding and acquisition of target vocabulary items. For instance, Webb (2008), found that the quality of the context is even more influential on the lexical encoding process than frequency of exposure. Moreover, Ramos and Dario (2015, p. 158) posited that "vocabulary learning largely depends on the context surrounding each word and the amount of attention that the learner places on both meaning and form". In line with this, the digital game tasks in the present study provided the participants with a rich and versatile context in several ways. First, the participants had the opportunity to experience the target words in two distinctive but relevant contexts, i.e., textual and audiovisual. The importance of this feature of digital games for vocabulary acquisition is tenable from a psycholinguistic perspective, namely, the "dual coding theory". In a situation where a concept can be stored in the mind both linguistically and visually is an outcome of the "Dual Encoding" process (Nation, 2001). The digital game-based context of this study allowed the possibility of encoding the target words into the participants' lexicon via two channels, namely the textual, and audiovisual channels, which, according to the dual coding theory, leads to effective acquisition. This type of acquisition is effective because, if various aspects of words are encoded in the mental lexicon, their retention will be easier (Rott & Williams, 2003). Thus, the digital game task context enhanced the acquisition of different dimensions and scopes of target words by enabling the encoding of various attributes of the words, in this case textual/audiovisual, into the mental lexicon, thereby facilitating their retention. Second, the provision of the game guide and implementation of textual enhancement techniques gave participants the opportunity of exposing to and of noticing both the form and meaning of the target words, a necessary condition for generating form-meaning links in vocabulary tasks (Ellis, 1994).

The second cognitive perspective, which can be advanced to explain the effectiveness in this study of the digital game task on the acquisition of the target words, is the strategies used. Vocabulary learning strategies are supported and recommended for efficient vocabulary acquisition (Schmitt & McCarthy, 1997; Nation, 2001; Schmitt, 2010) since they can "make learning easier, faster, more enjoyable, more self-directed, more effective, and more transferable to new situations" (Oxford, 1990, p. 8). In this study, the concurrent think-aloud data showed that the participants employed both cognitive and social strategies that suited their ongoing task, personal needs, contexts, and hence their attempts at solving problems. They tried to recall the definitions, requested feedback,

inferred meanings, applied their word-association knowledge, negotiated with their partners about the target words, and emphasized them by using enhancing techniques. The application of these strategies might have provided them with opportunities for processing the form, meaning and connections of the target words. It probably also enhanced their noticing of the dimensions and scopes of the target words, which is the starting point of language learning (Gass, 1998, as cited in Cross, 2002, p.3; Truscott, 1998). Thus, the encouragement to apply vocabulary learning strategies offered by the digital game tasks may have enhanced a deeper acquisition of different aspects of the target words as well as their forms and meanings.

In addition to context and strategies, the digital game might have cognitively enhanced the acquisition of dimensions and scopes of the target words via other cognitive factors. For instance, the task persuaded participants to use the target words both receptively for understanding the textual and in-game commands and productively for negotiating with their partners for planning or hypothesizing their future actions. In this case, Gass (1999) argued that learning both productively and receptively elevates learning aspects of vocabulary. Next, the target words were considered as tools for reaching a specific aim, that is to say, the digital game task drove the participants to interact with the target words to progress in the game. This created the feeling of relevance between the target words and the task, a factor which guarantees effective vocabulary acquisition (Laufer, 2001). Finally, due to the relation of the target words to the tasks, both the task and the textual/audiovisual context were meaningfully linked, further supporting long-term retention of the vocabulary items (Ramos & Dario, 2015).

Overall, it can be put that the interaction with the digital game tasks allowed the participants to experience contexts that presented the target word forms through both linguistic and audiovisual channels, and that this facilitated the encoding process of the target words by generating strong links between the various aspects of form and meaning of the target words. Moreover, they were encouraged to employ various strategies to process the target words. Therefore, due to rich cognitive support from the digital game, the newly generated stimuli might have been encoded richly into participants' lexicon and different dimensions and scopes of the vocabulary items might have been acquired successfully.

The success of digital game tasks in assisting in the acquisition of target words and their related aspects can also be discussed from a motivational perspective.

> Motivation has been widely embraced by both practitioners and researchers as a critical determinant of success in language learning, and this belief is strongly supported by a wide range of studies on L2 motivation … hence, it is logical to assume that motivation also facilitates vocabulary learning (Tseng & Schmitt, 2008, p. 385)

Motivating and engaging are qualities that have been attributed to digital games (Figg & Jaipal, 2009). Garris, Ahlers, and Driskell (2002) indicate that desirable learning is the outcome of motivation, which in digital games is increased by

elements like control and confidence, it. Moreover, "interactivity, rules, goals, challenge, risk, fantasy, curiosity, and control" (Pivec, Dziabenko, & Schinnerl, 2003, p. 220) increase motivation in a learner to a level that keeps the learner on task and boosts the desire to continue. Also, digital games contain elements such as challenge, fantasy and curiosity that enhance internal motivation (Dickey, 2006 as cited in Dondlinger, 2007). Among the different genres of digital games, the adventure genre has been found to contain the richest motivation-enhancing elements, especially challenge (Chen, Tseng, & Hsiao, 2018), and have been discussed as the most effective genre for vocabulary acquisition (Laveborn, 2009; Vahdat & Rasti Behbahani, 2013; Chen & Yang, 2013).

The sources of motivation in digital games have also be considered as predictive factors that can measure language learning (Allen, 2014 as cited in Ebrahimzadeh & Alavi, 2016), including vocabulary acquisition. Ebrahimzadeh and Alavi (2016) found that challenge, feedback and immersion were correlated with successful vocabulary acquisition through digital game tasks (Ebrahimzadeh & Alavi, 2016). They concluded that the provision of appropriate levels of motivation-enhancing elements increases the opportunity for vocabulary acquisition in digital game tasks. Moreover, a recent study also found that more interactivity-prone digital games can increase vocabulary acquisition better than less interactivity-prone games (Zhonggen, 2018); hence, a high level of interactivity can have a reverse result on vocabulary recall (deHaan, Reed, & Kuwada, 2010). Accordingly, I can infer that also the adventure digital game employed in this study might have provided proper levels of motivational-rising sources, especially, challenge, immersion, interactivity, and feedback, for the participants. Therefore, the multidimensionally motivating tasks might have sustained the participants in task and facilitated acquisition of the target words and their aspects. Otherwise, if the contributing factors in motivation, or the sources, failed to support the participants' emotion in order to sustain at the optimal level of motivation, they neither continued the task nor acquired the target words as well as their dimensions and scopes successfully (Kiili, 2005). The validity of this idea can be confirmed by referring to the exit-interview data where the participants told that elements such as story and gameplay, which are the umbrella terms for sources of motivation in the digital games (Ang & Zaphiris, 2006), were driving and motivating factors for them to sustain in the tasks; hence, sometimes, tasks were difficult and frustrating for them.

Various other factors, such as the role of feedback, noticing, controllability, repetition and word frequency, and instantiation, to name a few, can contribute to the effectiveness of vocabulary learning. However, as this was not the main purpose of this study, I limit this part to the above discussion.

6.2 Research Question 2: Which dimension and scope of word knowledge, either receptive (recall/recognition) or productive (recall/recognition), are acquired significantly better after completing digital game tasks in different levels of involvement load?

Unexpectedly, the results showed that the participants performed better in the productive post-test than receptive post-test. In other words, participants' productive knowledge of the target words was acquired better than their receptive knowledge after their interaction with the digital game tasks. Jasso (2012) found that a commercial digital game may assist in the acquisition of productive knowledge of words. However, he was testing domain-specific nouns, like hairstyling-related words such as brush, hair dryer etc. In a recent study, Sundqvist (2019) also found that frequency of gameplay, in the long run, could effectively assist in the acquisition of productive knowledge of vocabulary items.

As already mentioned, the precedence of productive knowledge acquisition in this study contradicts the previous findings indicating the precedence of receptive over productive knowledge (Morton, 1977 as cited in Barcroft, 2004; Meara, 1997; Nation, 1990, 2001; Schmitt, 2008, 2010a). However, there might be reasons for the superiority of productive over receptive learning in digital game tasks. To explain this phenomenon, I draw on an analogy between the current study and some earlier studies on vocabulary acquisition and in psychology. In the vocabulary learning literature, Mondria & Wiersma (2004, p. 82) state that "equivalence of type of learning and type of test […] yield[s] better results than non-equivalence of learning and testing". This premise is in accordance with the "transfer-appropriate processing" hypothesis in psychology, which "emphasizes that the value of particular acquisition activities must be defined relative to particular goals and purpose. Furthermore, assumptions about quality and durability of the resulting memory traces can only be determined relative to appropriateness of testing situations" (Morris, Bransford, & Franks, 1977, p. 528). Based on these premises, it is possible that the approach to vocabulary acquisition in the present digital game tasks, might have been more productive-oriented than receptive-oriented. With respect to the transfer-appropriate processing hypothesis, since the present participants performed better in the productive test, the acquisition and testing protocols might have been evoked processes that favored the acquisition of productive rather than receptive knowledge. Otherwise, the participants' performance in the receptive post-test can be expected to have been at least equal to their performance in productive post-test. It has also been found that retention loss is greater for receptive than productive knowledge in delayed tests (Mondria & Wiersma, 2004), a finding which also supports my argument on the productive-oriented nature of the digital game tasks.

However, how do I defend such a claim? I can answer this question by drawing another analogy between the nature of the task in this study and the

nature of the task in another study in the vocabulary acquisition literature. de la Fuente (2002) compared input-oriented tasks, in which negotiation is not necessary, to tasks that either require negotiation or negotiation and the production of novel utterances. She found that the input-only tasks were not effective enough in boosting the acquisition of vocabulary items. In turn, tasks encouraging negotiation plus production promoted productive knowledge better than receptive knowledge in vocabulary acquisition. In her study, the negotiated interaction plus production group were able to both elaborate and discuss the target words. In other words, they had the "opportunity to modify their own output and produce the target words [by] giving instructions …[and] asking questions about target words by paraphrasing or elaborating" (de la Fuente, 2002, p. 89). Clearly, the nature of her task and the nature of the task in my study are almost identical. In my study, the participants were performed the digital game tasks in pairs. In so doing, they had the opportunity to interact, negotiate, discuss, paraphrase, ask questions, elaborate, and produce the target words. In other words, experienced negotiated interaction plus production indirectly. The participants in this study, like those in de la Fuente's study, performed better in the productive than receptive post-test.

The reasons why the digital game task boosted acquisition of the productive aspect of the target words can also be discussed and explained. Ellis, Tanaka, & Yamakazi (1994, as cited in de la Fuente, 2002, p. 86) "suggested that negotiation may benefit productive acquisition of new words, provided that the students have the opportunity to use items they have begun to acquire and receive feedback from other speakers". The nature of the present task clearly meets the above-mentioned conditions of negotiation, using the items, and receiving feedback. The participants had to interact with each other in order to solve the problem in the digital game task and progress in the game. For instance, if the problem was the meaning of the target word, they would need to negotiate their understanding of the target words with their partner in generate new ideas for their future actions, hypotheses, and plans. Furthermore, they had to test their plans and hypotheses by playing the game. In so doing, they could receive feedback from either their partner or the digital game itself. Therefore, the participants' engagement in a task which provided both feedback and opportunities for using the item they had learned might have been one of the reasons for the primacy of productive learning.

It can also be speculated that the nature of the encoding process invoked processing of the new lexical items. Swain (1985) argued that production in language learning tasks triggers a different type of processing from input. Moreover, production concentrates the attention on the formal aspects of words rather than their semantic features. Thus, the target words are internalized such that the language learner can retrieve the form of the target word better than its meaning. Accordingly, in this study, it can be assumed that after the participants had recalled, inferred, or noticed the meaning of a target word, they might have needed to check it with their partner by pronouncing it. This might have prompted them with the to focus on form which in turn might have triggered

different and deeper processes in their minds for both code-breaking, or phonological analysis, and then decoding, or proper articulation (Swain, 1985; de la Fuente, 2002). Thus "it is obvious that […] discussion and negotiation, and multimode exposure to target vocabulary, [such as offered by a digital game], are all means of scaffolding and manipulating vocabulary that increased learners' use of target vocabulary" (Lee & Muncie, 2006, p. 312), leading to the effective acquisition of productive knowledge of the target vocabulary items.

Other possible reasons do not concern the nature of the task; instead they are relevant to the data collection procedure and the nature of the target words. Regarding the data collection procedure, a third speculation pertains to the order of the post-tests. Mondria and Wiersma (2004) posit that preceding receptive tests by productive tests might boost performance in in the former. In the present case, as I administered the receptive test first, this might have had learning effects on the participants and boosted their performance in the forthcoming productive tests. This possibility might be considered and tested in future studies.

On the nature of the vocabulary items, the final speculation concerns the issue of decay. It is often claimed that decay is greater in receptive than productive knowledge. For example, Mondria and Wiersma (2004) reported that Griffin and Herley (1996), in their experiment with the productive and receptive nature of vocabulary acquisition, found that decay in receptive knowledge, acquired from receptive learning, is greater than decay in productive knowledge, acquired from productive learning. Accordingly, it could be speculated that due to the rapid decay in receptive knowledge, participants might have forgotten the target words after the three-week interval, leading to their poor performance in the receptive post-test. To confirm this idea, the immediate effect of a digital game task could be evaluated in future studies; however, this is, perhaps, unnecessary, since later retention of target words is a more desirable learning outcome.

Based on the above discussion and reasoning, it can be concluded that the superiority of productive recall/recognition knowledge over other types of vocabulary knowledge in a digital game-based vocabulary acquisition task may, in the first place, be due to task features, such as the gamers possibility to interact in pairs in performing the tasks, and the characteristics of the target words. A digital game also may have a secondary or accelerating role in this process. In other words, in this study, the digital game might have supplied the learners with contexts that invoked relevant and responsible processes for learning the features of words. If the invoked process leads to productive learning, as found in this study, the digital game will boost productive knowledge acquisition. Furthermore, if they are processes, the digital game elements will also accelerate receptive-oriented processes. This may be another reason why the digital game task was found effective for different aspects target word learning in this study, despite the specific challenges presented by the three different tasks.

6.3 Research Question 3: Does interacting with the digital game tasks, in different levels of involvement load, make significant differences in vocabulary acquisition?

The main intention of this study was to investigate the effect of digital game task-induced levels of involvement load on the acquisition of the target vocabulary items. Although discovering the optimal level of involvement load was not the main goal, it was found that the group C participants, whose task required a high involvement load (index 12), acquired and retained productive (recall/recognition) and receptive (recall/recognition) knowledge of the target words better than the participants in the other two groups. They also seemed to generate stronger form-meaning links than the other two groups, namely, B (index 9) and A (index 7). Therefore, these findings partially support the premise of the involvement load hypothesis and support the part of the literature that advocates the effectiveness and significance of the hypothesis in vocabulary acquisition (Jing & Jianbin, 2009; Kim, 2011; Huang, Wilson, & Eslami, 2012; Xie, Zou, Wang, & Wong, 2017; Zou, 2017).

However, with respect to the performance of group B, these findings also draw our attention to that part of the literature that questions the validity of the hypothesis (Folse, 2006; Martínez-Fernández, 2008; Yaqubi, Rayati, & Gorgi, 2010; Jahangiri & Alipour, 2014). The group B participants, whose task-induced moderate level of involvement load (index 9), showed poorer performance in the post-tests than the group A participants. this outcome is counter to the premise of the involvement load hypothesis according to which a higher level of involvement load is a predictor of better vocabulary acquisition (Hulstijn & Laufer, 2001). Opponents of the hypothesis (Folse, 2006; Keating, 2008; Kim, 2010; Zou, 2017) have, as proposed by Hulstijn and Laufer (2001), criticized the original method of quantifying task-induced involvement load, by referring to the lack of precision in measuring its components, such as need, search and evaluation, and especially the evaluation component. Thus, they assert that the predictive power of the involvement load hypothesis is not reliable. In this study, a modified and extended version of the task-induced involvement load index, namely Technique Feature Analysis (Webb & Nation, 2010), was used to improve the precision of the original indexing method (Hu & Nassaji, 2016; Chaharlang & Farvardin, 2018; Gohar, Rahmanian, & Soleimani, 2018). Notwithstanding, the results showed that the involvement load hypothesis remained inaccurate in predicting the effectiveness of vocabulary acquisition tasks. In this study, the tasks assigned to group A and C induced a higher level of involvement load than the tasks assigned to group B; however, it may be that lack of precision in indexing had misled me into believing that group B's tasks induced a moderate level of involvement load. The criticism of the inaccuracy of the indexing of the evaluation component of the involvement load hypothesis supports my reasoning. Therefore, it seems reasonable to argue that the performance of the group B participants was poor because the induced level of the evaluation

component, which is crucial for the initial learning of vocabulary (Kim, 2011), was lower in group B's tasks than in the tasks of groups A and C. However, this was not revealed owing to the inaccuracy of the indexing method. On the other hand, it can also be argued that the indexing was precise and that other factors, like the split-attention effect, might have influenced group B's performance. Hence, we cannot place all the responsibility at the door of the involvement load hypothesis or the indexing methods used, as it is difficult to encapsulate the multidimensional nature of vocabulary acquisition in a single framework. Many other factors can contribute to either facilitating or impeding vocabulary acquisition. Thus, the poor performance of the group B participants and superior performance of the group A and C participants can also be explained by a number of other factors.

Poor performance of group B can be explained, for example, by such contributory factors as context, the mental processes the tasks involved, and the strategies the participants used. In regards with the role of context, previous studies indicate the superiority of meaning inferred glosses over the meaning given glosses (Rott, Williams, & Cameron, 2002; Nassaji, 2003; Rott, 2005). However, Hulstijn, Hollander, and Greidanus (1996) found that contextual clues are defining factors in the success of any inferring activities. From this point of view, the group B task was an inferring activity and the group A task, a meaning-giving activity. Accordingly, the poor performance of the group B participants might be due to the insufficient provision of contextual clues. Although the participants were supplied with a game guide to avoid them from missing in the game and gave them the chance of exposure to the target words, the guide was not contextually rich. To make it easily understandable, the game guide was composed as an instruction manual. Thus, it did not provide the group B participants with many clues to direct their inferring activity. Moreover, it encouraged form processing before meaning processing, which has been reported to be a negative factor in the acquisition of target words and its aspects (Ellis & He, 1999). Thus, they had no alternative but either to interact with the digital game context or to consult with their partners. Thus, although the context of the digital game and the game guide were supportive enough to help them learn target words, they were not so effective that could guarantee the acquisition of the target words as rich as the other two groups, i.e., A and C.

Mental processes and strategy choices are other possible contributory factors. As the qualitive data analysis revealed, the group B task encouraged the participants to employ inferencing techniques. Inferring from context invokes distinctive processes such as decision-making generators, or processes for deciding which definition to use, and evaluators or hypothesis testing, to test the selected definitions (Nassaji, 2003). Although these processes have been found to support vocabulary acquisition (Hulstijn, Hollander, & Greidanus, 1996), they do not explicitly initiate form-meaning links. Furthermore, due to multiple alternatives in meaning-inferred glosses, these processes might have been unnecessary and led to confusion (Martínez-Fernández, 2008; Bao, 2015). Thus, as the think-aloud data analysis seems to suggest, the participants may have been

puzzled by the many alternatives and floundered in their decision-making and hypothesis-testing. As a result, they did not manage to invoke the form-meaning link. Furthermore, the existence of possibilities such as trial-and-error, and cursor-hovering-name-reveal, in the digital game context, might have reduced any chances of form-meaning link formation because the constant clicking or hovering of the cursor over objects in the game could have assisted the participants to overcome the difficulties they encountered in each section of the digital game. Consequently, the participants did not feel the need to engage inferring meanings and solving problems related to the target words. Thus, the target words may not have been processed richly enough to be effectively retained by the group B participants.

Next, the reasons for the superiority of the group C participants can be speculated. The first possible explanation relates to the effect of frequency of exposure. Frequency of exposure, or repetition, is always mentioned as a facilitating factor in the process of either incidental or intentional vocabulary acquisition (Schmitt & McCarthy, 1997; Nation, 2001; Folse, 2006; Schmitt, 2008, 2010; Webb, 2007b; Bao, 2015). In this study, the group C participants had a greater chance of exposure to the target words than the other participants. They were exposed to them once before their main task in their target word list and sample sentences, and once more in the game guide. The nature of exposure was also spaced repetition, or exposure to the same word after a specific time interval. In the vocabulary acquisition literature, spaced repetition has been reported to be more effective than frequent repetition (Nation, 2001; Webb, 2007b). The higher frequency of exposure might also have increased their level of awareness, which in combination with an appropriate level of involvement load could promote vocabulary acquisition (Martínez-Fernández, 2008). Thus, the participants had more opportunities to process the target words and to become aware of their different aspects and features. This in turn could lead to deeper memory traces (Craik & Lockhart, 1979; Ellis & Beaton, 1993a), generate stronger form-meaning connections, and help the learners acquire the target words much more effectively.

Finally, the superior performance of group C can also be attributed to the role of pre-teaching. Laufer (2006) found that focus on forms, that is, in the case of vocabulary acquisition, receiving individual target words and their definitions in a list before a main task produces more desirable vocabulary acquisition outcomes than a focus on form or reading comprehension. Here, the group C participants were pre-taught the target words via a word list. They had, therefore, the opportunity to focus on both forms and form in their task. Given Laufer's findings, the group C's superior performance was not surprising. Group A may also have benefitted from a focus on forms as they received the target words in marginal glosses. However, why the group C outperformed group A in the post-tests can be explained by the presence of repetition, which may have deepened the process of encoding new lexical items and helped in creating deeper memory traces.

In the end, with respect to the involvement load hypothesis, the poor performance of group B and superiority of group C, it can be concluded that, if

we consider the indexing method as accurate, the task of group B, , failed to induce the expected moderate level of involvement load due to contributory factors such as context, mental processes, features of the digital game, and strategies. Consequently. As a result, the premise of the involvement load hypothesis was not met, in this study, for by the task designed for group B. However, group C's digital game task induced a higher level of involvement load and achieved its aims by other means such as pre-teaching, focus on forms, frequency of exposure, and spaced repetition. To conclude, involvement load is an important factor in acquiring the target words in digital game-based as well as traditional tasks. However, it lacks adequate predictive power for evaluating the success rate of vocabulary acquisition, in either digital game-based or traditional, tasks. Therefore, I believe that, although involvement load hypothesis is an important factor in vocabulary acquisition, it must be approached with care because it lacks adequate predictive power.

6.4 Conclusion

This study on the effectiveness of digital game tasks for vocabulary acquisition revealed that in spite of the variation in the constructive elements of the tasks, in this case cognitive elements, the chosen digital game supported vocabulary acquisition processes. Although the tasks were performed at different levels of involvement load, the target words were effectively acquired, if not wholly via the expected route, as the productive knowledge was better acquired than receptive knowledge. Thus, this study partially supported the involvement load hypothesis, showing that the level of involvement load influenced the outcome of the digital game-based vocabulary acquisition tasks.

A general conclusion that can be drawn from this study is that multimedia contexts, especially the digital game-based context, cannot be treated in the same way as the traditional pencil and paper contexts with regard to vocabulary acquisition. While task involvement load is a defining factor in traditional vocabulary acquisition contexts, it does not, owing to various factors such as task design, cognitive load, target word characteristics, mental processes, and strategy choices, play a major role in vocabulary acquisition in multimedia contexts. . Therefore, despite relatively high levels of involvement load, distinct differences can be found in participants' performance in digital game-based vocabulary acquisition tasks. Moreover, the use of inferencing techniques is not encouraged in the digital game-based vocabulary acquisition tasks even though it induces high level of involvement load and it is considered effective in traditional contexts. The reason for this is that it seems to lead to confusion rather than acquisition. Or, if inferencing is encouraged, it should be implemented carefully, meaning the provision of enough contextual clues and selecting digital games that support learning about the facilitating options Moreover, applying pre-teaching techniques and introducing the target words, their meanings and their uses plus higher exposures to the target words are highly recommended in

digital game-based vocabulary acquisition tasks. This is unlike traditional contexts, where inferencing has been highly encouraged for incidental vocabulary learning (Hulstijn, Hollander, & Greidanus, 1996; Rott, 2005). It seems that pre-teaching techniques can boost the effectiveness of digital game tasks and result in more desirable outcomes. If pre-teaching is not possible, meaning-given glosses are highly preferred to meaning-inferred glosses in vocabulary acquisition digital game tasks as they increase the possibility of precise form-meaning link generation and prevent confusion.

Finally, the most noteworthy findings of this study, such as the better acquisition of productive knowledge and the poor performance of the group B participants in the post-tests, are emphatic reminders of the multidimensional nature of vocabulary acquisition, which that makes vocabulary more difficult to acquire than other components of language. Furthermore, these findings indicate that the multidimensional nature of vocabulary acquisition cannot be explained from a single point of view, such as, in the present instance, the involvement load hypothesis. Several other factors, such as context, frequency of exposure, cognitive load and decay, to name a few, that can alter the expected outcomes of every vocabulary acquisition task and the predictions of vocabulary acquisition hypotheses and theories. The effectiveness of digital games in vocabulary acquisition supports the multidimensional nature of vocabulary acquisition since the digital game context offers multiple factors that promote learning processes and facilitate acquisition. In this study, irrespective of the differences in task types and constructs between different groups of participants, the digital game was boosted the acquisition of vocabulary items.

6.5 Limitations of the study and suggestion for further research

Like every other study in the humanities, this study has its limitations. The main limitation concerns the target words. In the present instance, the target words were selected only from one lexical class, that is, nouns. More specifically, they were all concrete nouns. Thus, caution is advised in generalizing the findings of this study to other lexical classes and to abstract nouns. The influence of lexical category is a possible topic for future studies.

The lack of a control group is another potential limitation of this study. Although a control group might have increased the validity of the quantitative results, I preferred not to include one because, as the topic of this thesis suggests, I was *investigating* rather than measuring or evaluating the effect. To investigate the effect, I collected the qualitative data that allowed me to track online mental process and to observe how they were modified by the digital game tasks. However, this study could be replicated, including a control group, and the results compared to those of this study.

As discussed above, the superiority of productive over receptive target word knowledge might have been due to the prior administration of the receptive post-test. It seems that the order of administration could have

positively influenced participants' productive knowledge. This study could thus be replicated with the post-test administered in the reverse order. The results should then be compared to the present to reveal if the higher scores on productive knowledge was the outcome of either the digital game vocabulary acquisition task or the order of administration of the post-tests.

Finally, the low number of target words and participants hinder the generalization of the present results. Tis study comprised only 20 target words and 30 participants. Increasing the number of participants and target words in future studies would be very helpful in either confirming or rejecting the present outcomes and what they add to our knowledge on the digital game-based tasks and vocabulary acquisition.

6.6 Implications of this study

Previous digital game and vocabulary acquisition studies have mostly concluded that digital games are better used in a complementary role as an extramural vocabulary acquisition activity (Sundqvist & Sylvén, 2012). However, this study, by simulating a commonplace language classroom task and integrating elements such as text and peer interaction, endeavored to push digital game vocabulary acquisition tasks a step closer to their integration in the language learning classroom by including them as a classroom activity. In this regard, the present findings of this study may benefit language teachers and researchers.

In light of the outcomes of this study, it is important that language teachers are aware that integrating digital games into language courses, while possible, must be done with care. They should know that the pre-teaching of the target words, provision of a game guide offering the possibility of peer interaction, and that implementing a high level of involvement load are prerequisites of both quality and quantity in target word acquisition and retention. Furthermore, they must be aware that the genre of the digital game is also important. The adventure genre is highly recommended over other genres for boosting vocabulary acquisition due to its functional mixture of gameplay and story. Thus, teachers can ensure that both the motivational and mental requirements for high quality task performance and efficient vocabulary acquisition are met. On the other hand, teachers should perhaps avoid inferencing techniques in digital game tasks as these strategies not only do not assist vocabulary acquisition but also confuse language learners.

Finally, these findings also have implications for researchers. First, to control for unintended factors, researchers should consider these findings in their future digital game-based vocabulary acquisition studies. Moreover, these results may help reveal novel aspects and features that help boost the impact of digital game tasks and game designers in producing digital games that serve the purposes of both education and entertainment.

REFERENCES

Ahmad, J. (2011). Intentional vs. incidental vocabulary learning. *Interdisciplinary Journal of Contemporary Research in Business, 3*(5), 67-75.

Alcon, E. (2007). Incidental focus on form, noticing and vocabulary learning in the EFL classroom. *International Journal of English Studies, 7*(2), 41-60.

Alexiou, A., & Schippers, M. C. (2018). Digital game elements, user experience and learning: A conceptual framework. *Education and Information Technologies*, *23*(6), 2545-2567.

Alias, N., Rosman, F., Rahman, M. N. A., & Dewitt, D. (2015). The potential of video game in Malay language learning for foreign students in a public higher education institution. *Procedia-Social and Behavioral Sciences*, *176*, 1020-1027.

Anderson, T. A., Reynolds, B. L., Yeh, X., & Huang, G. (2008). Video games in the English as a foreign language classroom. Paper presented at the *2008 Second IEEE International Conference on Digital Game and Intelligent Toy Enhanced Learning*, 188-192.

Ang, C. S., & Zaphiris, P. (2006). Developing enjoyable second language learning software tools: A computer game paradigm. *User-centered computer aided language learning* (pp. 1-21) Hershey, PA: IGI Global.

Ang, C. S., & Zaphiris, P. (2008). Computer games and language learning. *Handbook of Research on Instructional Systems and Technology*. Hershey, PA: IGI Global.

Arjoranta, J. (2014). Game definition: A Wittgenstein approach. *The International Journal of Computer Game Research, 14* (1). Retrieved March, 18th, 2019 at http://gamestudies.org/1401/articles/arjoranta

Bakar, N. A., & Nosratirad, E. (2013). Sustaining vocabulary acquisition through computer game: A case study. *Asian Social Science*, *9*(5), p235.

Bao, G. (2015). Task type effects on English as a foreign language learners' acquisition of receptive and productive vocabulary knowledge. *System*, *53*, 84-95.

Barcroft, J. (2004). Second language vocabulary acquisition: A lexical input processing approach. *Foreign Language Annals, 37*(2), 200-208.

Becker, K. (2008). Video Game Pedagogy. In *Games: Purpose and potential in education* (pp. 73-125). Springer, Boston, MA.

Beglar, D. (2010). A Rasch-based validation of the Vocabulary Size Test. *Language testing*, *27*(1), 101-118.

Bloomfield, L. (1933). *Language*. New York: Henry Holt.

Boers, F., Eyckmans, J., Kappel, J., Stengers, H., & Demecheleer, M. (2006). Formulaic sequences and perceived oral proficiency: Putting a lexical approach to the test. *Language Teaching Research, 10*(3), 245-261.

Boers, F., Warren, P., Grimshaw, G., & Siyanova-Chanturia, A. (2017). On the benefits of multimodal annotations for vocabulary uptake from reading. *Computer Assisted Language Learning*, *30*(7), 709-725.

Bowers, P. N., & Kirby, J. R. (2010). Effects of morphological instruction on vocabulary acquisition. *Reading and Writing*, *23*(5), 515-537.

Bowles, M. A. (2010). The think-aloud controversy in second language research. New York: Routledge.

Brown, J. D. (1988). *Understanding research in second language learning: A teacher's guide to statistics and research design*. Cambridge: Cambridge University Press.

Brown, T. S., & Perry Jr, F. L. (1991). A comparison of three learning strategies for ESL vocabulary acquisition. *TESOL Quarterly*, *25*(4), 655-670.

Carter, R. (1998). *Vocabulary: Applied linguistics perspective*. London: Routledge.

Chaharlang, N., & Farvardin, M. T. (2018). Predictive Power of Involvement Load Hypothesis and Technique Feature Analysis across L2 Vocabulary Learning Tasks. *International Journal of Foreign Language Teaching and Research*, *6*(24), 127-141.

Chang, A., & Hu, H, M. (2018). Learning Vocabulary through Extensive Reading: Word Frequency Levels and L2 Learners' Vocabulary Knowledge Level. *TESL-EJ*, *22*(1), 1 - 20.

Charters, E. (2003). The use of think-aloud methods in qualitative research an introduction to think-aloud methods. *Brock Education: A Journal of Educational Research and Practice, 12*(2). 68-82.

Chen, H. J. H., & Hsu, H. L. (2019). The impact of a serious game on vocabulary and content learning. *Computer Assisted Language Learning*, 1-22.

Chen, H. J. H., & Yang, T. Y. C. (2013). The impact of adventure video games on foreign language learning and the perceptions of learners. *Interactive Learning Environments*, *21*(2), 129-141.

Chen, M. H., Tseng, W. T., & Hsiao, T. Y. (2018). The effectiveness of digital game-based vocabulary learning: A framework-based view of meta-analysis. *British Journal of Educational Technology*, *49*(1), 69-77.

Chian-Wen, K. A. O. (2014). The Effects of Digital Game-based Learning Task in English as a Foreign Language Contexts: A Meta-analysis. *Education Journal*, *42*(2), 113-141.

Chiu, Y. H., Kao, C. W., & Reynolds, B. L. (2012). The relative effectiveness of digital game-based learning types in English as a foreign language setting: A meta-analysis. *British Journal of Educational Technology*, *43*(4), E104-E107.

Cohen, A. D. (1987). Using verbal reports in research on language learning. In C. Faerch & G. Kasper (Eds.), *Multilingual matters, No. 30. Introspection in second language research* (pp. 82-95). Clevedon, England: Multilingual Matters.

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of Processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior, 11*(6), 671-684.

Craik, F. I. M., & Tulving, E. (1975). Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology, 104*(3), 268-294.

Cross, J. (2002). Noticing in SLA: Is it a valid concept. *TESL-EJ*, *6*(3), 1-9.

Czikszentmihalyi, M. (1990). *Flow: The psychology of optimal experience*. New York: Harper & Row.

De Freitas, S., & Maharg, P. (Eds.). (2011). *Digital games and learning*. London: Bloomsbury Publishing.

De La Fuente, M. J. (2002). Negotiation and oral acquisition of L2 vocabulary: The roles of input and output in the receptive and productive acquisition of words. *Studies in second language acquisition*, *24*(1), 81-112.

DeHaan, J. W. (2005). Acquisition of Japanese as a foreign language through a baseball video game. *Foreign Language Annals*, *38*(2), 278-282.

DeHaan, J. W. (2008). *Video games and second language acquisition: The effect of interactivity with a rhythm video game on second language vocabulary recall, cognitive load, and telepresence* (Doctoral dissertation, New York University, New York).

Dehaan, J., Reed, W. M., & Kuwada, K. (2010). The effect of interactivity with a music video game on second language vocabulary recall. *About Language Learning & Technology*, *14*(2), *74-94*.

Dondlinger, M. J. (2007). Educational video game design: A review of the literature. *Journal of applied educational technology*, *4*(1), 21-31.

Dörnyei, Z. (2007). *Research methods in applied linguistics: Quantitative, qualitative, and mixed methodologies*. Oxford: Oxford University Press.

Ebrahimzadeh, M., & Alavi, S. (2016). Motivating EFL students: E-learning enjoyment as a predictor of vocabulary learning through digital video games. *Cogent Education*, *3*(1), 1-14.

Elgort, I. (2012). Effects of L1 definitions and cognate status of test items on the Vocabulary Size Test. *Language Testing*, *30*(2), 253-272.

Ellis, N. C. (1995). The psychology of foreign language vocabulary acquisition: Implication for CALL. *International Journal of Computer Assisted Language Learning (CALL), 8*(2-3), 103-128.

Ellis, N. C. (1997). Vocabulary acquisition: Word structure, collocation, word-class, and meaning. in N. Schmitt & M. McCarthy (eds.), *Vocabulary: description, acquisition and pedagogy* (pp. 123-139). Cambridge: Cambridge University Press.

Ellis, N. C., & Beaton, A. (1993a). Psycholinguistic determinants of foreign language vocabulary learning. *Language Learning, 43*(4), 559-617.

Ellis, N. C., & Sinclair, S. G. (1996). Working memory in the acquisition of vocabulary and syntax: Putting language in good order. *The Quarterly Journal of Experimental Psychology Section A, 49*(1), 234-250.

Ellis, N., & Beaton, A. (1993b). Factors affecting the learning of foreign language vocabulary: Imagery keyword mediators and phonological short-term memory. *The Quarter Journal of Experimental Psychology, 46*(3), 533-558.

Ellis, R. (1994). Factors in the incidental acquisition of second language vocabulary from oral input: A review essay. *Applied Language Learning*, *5*(1), 1-32.

Ellis, R. (2009). Implicit and explicit learning, knowledge and instruction. *Implicit and explicit knowledge in second language learning, testing and teaching*, *42*, 3-25.

Ellis, R. (2017). Task-based language teaching. In *The Routledge handbook of instructed second language acquisition* (pp. 124-141). New York: Routledge.

Ellis, R., & He, X. (1999). The roles of modified input and output in the incidental acquisition of word meanings. *Studies in second language acquisition*, *21*(2), 285-301.

Elo, S., & Kyngäs, H. (2008). The qualitative content analysis process. *Journal of advanced nursing*, *62*(1), 107-115.

Ericsson, K. A., & Simon, H. A. (1980). Verbal reports as data. *Psychological review*, *87*(3), 215.

Ericsson, K. A., & Simon, H. A. (1987). Verbal reports on thinking. In C. Faerch & G. Kasper (Eds.), *Multilingual matters, No. 30. Introspection in second language research* (pp. 24-53). Clevedon, England: Multilingual Matters.

Ericsson, K. A., & Simon, H. A. (1993). *Protocol analysis: Verbal reports as data*. Cambridge, MA: MIT press.

Ermi, L., Heliö, S., & Mäyrä, F. (2004). Pelien voima ja pelaamisen hallinta. *Lapset ja nuoret pelikulttuurien toimijoina. Tampereen yliopiston hypermedialaboratorion verkkojulkaisuja*, *6*.

Figg, C., & Jaipal, K. (2009). Using video games in science instruction: Pedagogical, social, and concept-related aspects. *Journal of Canadian Journal of Science, Mathematics and Technology Education, 9(2)*, 117-134.

Fillmore, C. J., Kay, P., & O'connor, M. C. (1988). Regularity and idiomaticity in grammatical constructions: The case of let alone. *Language*, 501-538.

Folse, K. S. (2006). The effect of type of written exercise on L2 vocabulary retention. *TESOL quarterly*, *40*(2), 273-293.

Fotouhi-Ghazvini, F., Earnshaw, R. A., Robison, D., & Excell, P. S. (2009). The MOBO City: A Mobile Game Package for Technical Language Learning. *International Journal of Interactive Mobile Technologies*, *3*(2), 19-24.

Fraser, C. A. (1999). Lexical processing strategy use and vocabulary learning through reading. *Studies in second language acquisition*, *21*(2), 225-241.

Fries, C. C., & Traver, A. A. (1940). English word lists; a study of their adaptability for instruction. American Council on Education, Michigan: University of Michigan.

Garris, R., Ahlers, R., & Driskell, J. E. (2002). Games, motivation, and learning: A research and practice model. *Simulation & gaming*, *33*(4), 441-467.

Gass, S. (1999). Discussion: Incidental vocabulary learning. *Studies in second language acquisition*, *21*(2), 319-333.

Gathercole, S. E., & Baddeley, A. D. (1990). The role of phonological memory in vocabulary acquisition: A study of young children learning new names. *British Journal of Psychology*, *81*(4), 439-454.

Gee, J. P. (2003). *What video games have to teach us about learning and literacy*. Palgrave Macmillan.

Gohar, M. J., Rahmanian, M., & Soleimani, H. (2018). Technique feature analysis or involvement load hypothesis: estimating their predictive power in vocabulary learning. *Journal of psycholinguistic research*, *47*(4), 859-869.

Gu, P. Y. (2003). Vocabulary learning in a second language: Person, task, context and strategies. *TESL-EJ*, *7*(2), 1-25.

Gu, Y., & Johnson, R. K. (1996). Vocabulary learning strategies and language learning outcomes. *Language learning*, *46*(4), 643-679.

Hainey, T., Connolly, T. M., Boyle, E. A., Wilson, A., & Razak, A. (2016). A systematic literature review of games-based learning empirical evidence in primary education. *Computers & Education*, *102*, 202-223.

Hall, J. L. (2015). *A guide to doing statistics in second language research using SPSS and R*. Routledge: New York.

Hamari, J., Koivisto, J., & Sarsa, H. (2014, January). Does gamification work?-A literature review of empirical studies on gamification. In *2014 47th Hawaii international conference on system sciences (HICSS)* (pp. 3025-3034). IEEE.

Hamari, J., Shernoff, D. J., Rowe, E., Coller, B., Asbell-Clarke, J., & Edwards, T. (2016). Challenging games help students learn: An empirical study on engagement, flow and immersion in game-based learning. *Computers in human behavior*, *54*, 170-179.

Haunted Hotel: Death Sentence & Eternity. (2017a, October 21). Retrieved from http://www.esrb.org/ratings/Synopsis.aspx?Certificate=34828&Title=Haunted+Hotel%3a+Death+Sentence+%26+Eternity

Haunted Hotel: Death Sentence & Eternity. (2017b, October 21). Retrieved from http://www.bigfishgames.com/games/8489/haunted-hotel-death-sentence/?pc

Hong, X. (2010). Review of Effects of Glosses on Incidental Vocabulary Learning and Reading Comprehension. *Chinese Journal of Applied Linguistics (Foreign Language Teaching & Research Press)*, *33*(1).

Hu, H. C. M., & Nassaji, H. (2016). Effective vocabulary learning tasks: involvement load hypothesis versus technique feature analysis. *System*, *56*, 28-39.

Huang, S., Willson, V., & Eslami, Z. (2012). The effects of task involvement load on L2 incidental vocabulary learning: A meta-analytic study. *The Modern Language Journal*, *96*(4), 544-557.

Huckin, T., & Coady, J. (1999). Incidental vocabulary acquisition in a second language: A review. *Studies in Second Language Acquisition, 21*(2), 181-193.

Hulstijn, J. H. (2001). Intentional and incidental second-language vocabulary learning: A reappraisal of elaboration, rehearsal and automaticity. In P. Robinson (Ed.), *Cognition and Second Language Instruction* (pp.258-286). Cambridge: Cambridge University Press.

Hulstijn, J. H., & Laufer, B. (2001). Some empirical evidence for the involvement load hypothesis in vocabulary acquisition. *Language Learning, 51*(3), 539-558.

Hulstijn, J. H., Hollander, M., & Greidanus, T. (1996). Incidental vocabulary learning by advanced foreign language students: The influence of

marginal glosses, dictionary use, and reoccurrence of unknown words. *The modern language journal*, *80*(3), 327-339.

Hung, H. T., Yang, J. C., Hwang, G. J., Chu, H. C., & Wang, C. C. (2018). A scoping review of research on digital game-based language learning. *Computers & Education*, *126*, 89-104.

Hung, K. H. (2011). *The design and development of an education-designed massively multiplayer online role-playing game (edd mmorpg) for young Taiwanese mandarin-speaking learners learning English vocabulary words* (Doctoral dissertation, Teachers College, Columbia University).

Hunt, A. (2009). Dictionaries and vocabulary learning: the roles of L1 and L2 information. *Journal of foreign language studies*, (1), 13-25.

Hwang, G. J., & Wang, S. Y. (2016). Single loop or double loop learning: English vocabulary learning performance and behavior of students in situated computer games with different guiding strategies. *Computers & Education*, *102*, 188-201.

Jackson, G. T., Dempsey, K. B., & McNamara, D. S. (2012). Game-based Practice in a Reading Strategy Tutoring System: Showdown in iSTART-ME. *Digital Games in Language Learning and Teaching (New Language Learning and Teaching Environments)*, 115-138.

Jahangiri, K., & Abilipour, I. (2014). Effects of collaboration and exercise type on incidental vocabulary learning: Evidence against involvement load hypothesis. *Procedia-Social and Behavioral Sciences*, *98*, 704-712.

Janebi Enayat, M., & Haghighatpasand, M. (2019). Exploiting adventure video games for second language vocabulary recall: a mixed-methods study. *Innovation in Language Learning and Teaching*, *13*(1), 61-75.

Jasso, P. R. (2012). A non-academic computer video game: Its effect on vocabulary acquisition in the EFL classroom (Published master's thesis). Syracuse university, New York.

Jenkins, J. R., Stein, M. L., & Wysocki, K. (1984). Learning vocabulary through reading. *American Educational Research Journal*, *21*(4), 767-787.

Jing, L., & Jianbin, H. (2009). An empirical study of the involvement load hypothesis in incidental vocabulary acquisition in EFL listening. *Polyglossia, 16*, 1-11.

Joe, A. (1998). What effects do text-based tasks promoting generation have on incidental vocabulary acquisition? *Applied Linguistics 19*(3), 357-377.

Johnson, W. L. (2007). Serious use of a serious game for language learning. *Frontiers in Artificial Intelligence and Applications*, *158*, 67.

Johnson, W. L., Vilhjálmsson, H. H., & Marsella, S. (2005, May). Serious games for language learning: How much game, how much AI?. In *AIED* (Vol. 125, pp. 306-313).

Kang, S. H. (1995). The effects of a context-embedded approach to second-language vocabulary learning. *System*, *23*(1), 43-55.

Katalin, E. (2000). Please keep talking": The 'think-aloud' method in second language reading research. *NovELTy VII*, *3*.

Keating, G. D. (2008). Task effectiveness and word learning in a second language: The involvement load hypothesis on trial. *Language Teaching Research*, *12*(3), 365-386.

Kiili, K. (2005). Digital game-based learning: Towards an experiential gaming model. *The Internet and higher education*, *8*(1), 13-24.

Kim, Y. (2011). The role of task-induced involvement and learner proficiency in L2 vocabulary acquisition. *Language learning*, *61*, 100-140.

Kirriemuir, J. (2002). *The relevance of video games and gaming consoles to the higher and further education learning experience*. Retrieved June 20, 2009, from http://www.jisc.ac.uk/index.cfm?name=techwatch_report_0201.

Kiss, G. R. (1973). Grammatical word classes: A learning process and its simulation. *Psychology of Learning and Motivation*, *7*, 1-41.

Kocić, A. (2008). The problem of synforms (similar lexical forms). *FACTA UNIVERSITATIS-Linguistics and Literature*, *6*(1), 51-59.

Krashen, S. (1989). We Acquire Vocabulary and Spelling by Reading: Additional Evidence for the Input Hypothesis. *The Modern Language Journal, 73*(4), 440-464. doi:10.2307/326879.

Krashen, S. (2009) *Principles and practice in second language acquisition*. New York: Pergamon. *Retrieved January, 12,* 2020.

Laufer, B. (1997). What's in a word that makes it hard or easy: Some intralexical factors that affect the learning of words. in N. Schmitt & M. McCarthy (eds.), *Vocabulary: description, acquisition and pedagogy* (pp. 140-155). Cambridge: Cambridge University Press.

Laufer, B. (2001). Reading, word-focused activities and incidental vocabulary acquisition in a second language. *Journal of Prospect, 16*(3), 44–54.

Laufer, B. (2006). Comparing focus on form and focus on forms in second-language vocabulary learning. *Canadian Modern Language Review*, *63*(1), 149-166.

Laufer, B. (2009). Second language vocabulary acquisition from language input and from form-focused activities. *Language Teaching*, *42*(3), 341-354.

Laufer, B., & Goldstein, Z. (2004). Testing vocabulary knowledge: Size, strength, and computer adaptiveness. *Language learning*, *54*(3), 399-436.

Laufer, B., & Hulstijn, J. (2001). Incidental vocabulary acquisition in the second language: The construct of task-induced involvement. *Applied Linguistics, 22*(1), 1-26.

Laveborn, J. (2009). Video Game Vocabulary: The effect of video games on Swedish learners‟ word comprehension. *Journal of Estetisk-filosofiska fakulteten,* 1-26.

Lee, S. H., & Muncie, J. (2006). From receptive to productive: Improving ESL learners' use of vocabulary in a post reading composition task. *Tesol Quarterly*, *40*(2), 295-320.

Leow, R. P., & Morgan-Short, K. (2004). To think aloud or not to think aloud: The issue of reactivity in SLA research methodology. *Studies in second language acquisition*, *26*(1), 35-57.

Levelt, W. J. M. (1992). Accessing words in speech production: Stages, processes and representations. *Cognition, 42*(1-3), 1-22.

Levelt, W. J., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and brain sciences*, *22*(1), 1-38.

Lin, C. C., & Huang, H. M. (2008). Meaning-inferred gloss and meaning-given gloss on incidental vocabulary learning. *Journal of National Taiwan Normal University: Humanities & Social Sciences*, *53*(2), 87-116.

Mackey, A., & Gass, S. M. (2015). *Second language research: Methodology and design*. Routledge: New York.

Martínez-Fernández, A. (2008). Revisiting the involvement load hypothesis: Awareness, type of task and type of item. In *Selected proceedings of the 2007 second language research forum* (pp. 210-228). Somerville, MA: Cascadilla Press.

Mayer, R. E., & Moreno, R. (2003). Nine ways to reduce cognitive load in multimedia learning. *Educational psychologist*, *38*(1), 43-52.

Meara, P. (1997). Towards a new approach to modelling vocabulary acquisition. in N. Schmitt & M. McCarthy (eds.), *Vocabulary: description, acquisition and pedagogy* (pp. 109-121). Cambridge: Cambridge University Press.

Milton, J. (2009). *Measuring second language vocabulary acquisition*. Bristol: Multilingual Matters.

Mondria, J. A., & Wiersma, B. (2004). Receptive, productive, and receptive+ productive L2 vocabulary learning: What difference does it make. *Vocabulary in a second language: Selection, acquisition, and testing*, *15*(1), 79-100.

Morgan, D. L. (1993). Qualitative content analysis: a guide to paths not taken. *Qualitative health research*, *3*(1), 112-121.

Morris, C. D., Bransford, J. D., & Franks, J. J. (1977). Levels of processing versus transfer appropriate processing. *Journal of Memory and Language*, *16*(5), 519-533.

Muhanna, W. (2012). Using online games for teaching English vocabulary for Jordanian students learning English as a foreign language. *Journal of College Teaching & Learning (TLC)*, *9*(3), 235-244.

Nassaji, H. (2003). L2 vocabulary learning from context: Strategies, knowledge sources, and their relationship with success in L2 lexical inferencing. *Tesol Quarterly*, *37*(4), 645-670.

Nation, I. S. P. (1990). *Teaching and learning vocabulary*. Boston, Mass.: Heinle & Heinle.

Nation, I. S. P. (2001). *Learning vocabulary in another language.* Cambridge: Cambridge University Press.

Nation, I. S. P. (2006). How large a vocabulary is needed for reading and listening? *Journal of the Canadian Modern Language Review, 63* (1), 59-82.

Nation, I. S. P., & Webb, S. A. (2010). *Researching and analyzing vocabulary*. Heinle: Cengage Learning Boston, MA.

Nation, I.S.P (2012). The vocabulary size test. Retrieved online at https://www.victoria.ac.nz/lals/about/staff/publications/paul-

nation/Vocabulary-Size-Test-information-and-specifications.pdf on 13 october 2017

Nation, I.S.P. & Beglar, D. (2007) A vocabulary size test. The Language Teacher, 31(7), 9-13.

Neville, D. O., Shelton, B. E., & McInnis, B. (2009). Cybertext redux: Using digital game-based learning to teach L2 vocabulary, reading, and culture. *Computer Assisted Language Learning*, *22*(5), 409-424.

Oxford, R. L. (1990). *Language Learning Strategies: What Every Teacher Should Know.* Boston: Heinle & Heinle.

Paribakht, T. S., & Wesche, M. (1997). Vocabulary enhancement activities and reading for meaning in second language vocabulary acquisition. In J. Coady & T. Huckin (Eds.), *Second language vocabulary acquisition: A rationale for pedagogy* (pp. 174–199). New York: Cambridge University Press.

Pasfield-Neofitou, S. (2014). Language learning and socialization opportunities in game worlds: Trends in first and second language research. *Language and Linguistics Compass*, *8*(7), 271-284.

Peters, E., & Webb, S. (2018). Incidental vocabulary acquisition through viewing L2 television and factors that affect learning. *Studies in Second Language Acquisition*, *40*(3), 551-577.

Peterson, M. (2010). Massively multiplayer online role-playing games as arenas for second language learning. *Computer Assisted Language Learning*, *23*(5), 429-439.

Piirainen-Marsh, A., & Tainio, L. (2009). Other-repetition as a resource for participation in the activity of playing a video game. *The Modern Language Journal*, *93*(2), 153-169.

Pivec, M., Dziabenko, O., & Schinnerl, I. (2003, July). Aspects of game-based learning. In *3rd International Conference on Knowledge Management, Graz, Austria* (pp. 216-225).

Plass, J. L., Moreno, R., & Brünken, R. (Eds.). (2010). *Cognitive load theory*. New York: Cambridge University Press.

Prensky, M. (2007). *Digital game-based learning*. Minnesota: Paragon House.

Qian, D. D. (1996). ESL vocabulary acquisition: Contextualization and decontextualization. *The Canadian Modern Language Review, 53*(1), 120–142.

Ramos, R., & Dario, F. (2015). Incidental vocabulary learning in second language acquisition: A literature review. *Profile Issues in TeachersProfessional Development*, *17*(1), 157-166.

Rankin, Y., Gold, R., & Gooch, B. (2006a). 3D role-playing games as language learning tools. In *Proceedings of EuroGraphics* (Vol. 25, No. 3, pp. 211-225).

Rankin, Y., Gold, R., & Gooch, B. (2006b). Evaluating interactive gaming as a language learning tool. In *Proceedings for ACM SIGGRAPH Conference*.

Rasti Behbahani, A. (2017, March). *The effect of video games on learning lexical word classes*, Presented at Nordic Educational Research Association Conference, Aalborg university, Denmark, March 2017.

Razak, A. A., Connolly, T., & Hainey, T. (2012). Teachers' views on the approach of digital games-based learning within the curriculum for excellence. *International Journal of Game-Based Learning (IJGBL)*, *2*(1), 33-51.

Read, J. (2000). *Assessing vocabulary*. New York: Cambridge University Press.

Read, J. (2004). 7. Research in teaching vocabulary. *Annual Review of Applied Linguistics*, *24*, 146-161.

Reinders, H., & Wattana, S. (2012). Talk to me! Games and students' willingness to communicate. *Digital games in language learning and teaching*, 156-188.

Reinhardt, J., & Sykes, J. M. (2012). Conceptualizing digital game-mediated L2 learning and pedagogy: Game-enhanced and game-based research and practice. *Digital Games in Language Learning and Teaching (New Language Learning and Teaching Environments)*, 32-49.

Reynolds, B. L. (2017). Evidence for the task-induced involvement construct in incidental vocabulary acquisition through digital gaming. *The Language Learning Journal*, *45*(4), 466-484.

Richards, J. C. (1976). The role of vocabulary teaching. *TESOL Quarterly, 10*(1), 77-89.

Ringbom, H. (1987). *The role of the first language in foreign language learning* (Vol. 34). Clevedon: Multilingual Matters Ltd.

Rott, S. (2005). Processing Glosses: A Qualitative Exploration of How Form-Meaning Connections Are Established and Strengthened. *Reading in a Foreign Language*, *17*(2), 95-124.

Rott, S., & Williams, j. (2003). Making form-meaning connections while reading: A qualitative analysis of word processing. *Reading in a Foreign Language*, *15*(1), 45.

Rott, S., Williams, J., & Cameron, R. (2002). The effect of multiple-choice L1 glosses and input-output cycles on lexical acquisition and retention. *Language Teaching Research*, *6*(3), 183-222.

Saravia, A. M. F. (1995). *Using the think aloud protocol to investigate vocabulary strategies used in two vocabulary tests* (Doctoral dissertation, Iowa State University Digital Repository).

Schlickum, M. K., Hedman, L., Enochsson, L., Kjellin, A., & Felländer-Tsai, L. (2009). Systematic video game training in surgical novices improves performance in virtual reality endoscopic surgical simulators: a prospective randomized study. *World journal of surgery*, *33*(11), 2360-2367.

Schmitt, N. (1997). Vocabulary learning strategies. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary: description, acquisition and pedagogy* (pp.199-228). Cambridge: Cambridge University Press.

Schmitt, N. (2007). Current perspectives on vocabulary teaching and learning. in J. Cummins & C. Davison (eds.), *The international handbook of English language teaching* (pp. 745-759). Springer.

Schmitt, N. (2008). Review Article: Instructed second language vocabulary learning. *Journal of Language Teaching Research, 12(3)*, 329-363.

Schmitt, N. (2010a). *Researching vocabulary: A vocabulary research manual.* Hampshire: Palgrave Macmilan.

Schmitt, N. (2010b). Key issues in teaching and learning vocabulary. in R. C. Beltran, C. A. Contesse, & M. M. T. Lopez (ed.), *Insights into Non-native Vocabulary Teaching and Learning* (pp. 28-40). UTP: Multilingual Matters.

Schmitt, N., & McCarthy, M. (1997). *Vocabulary: description, acquisition and pedagogy*. Cambridge: Cambridge University Press.

Scholz, K. (2017). Encouraging free play: Extramural digital game-based language learning as a complex adaptive system. *Calico journal*, *34*(1), 39-57.

Schrader, P. G., & McCreery, M. (2008). The acquisition of skill and expertise in massively multiplayer online games. *Educational Technology Research and Development*, *56*(5-6), 557-574.

Seliger, H. W., & Shohamy, E. G. (1989). *Second language research methods*. Oxford: Oxford University Press.

Shaffer, D., Squire, K., Halverson, R., & Gee, J. P. (2005). Video games and the future of learning. *Phi Delta Kappan*, *87*(2), 104-111.

Shahriarpour, N. (2014). On the Effect of Playing Digital Games on Iranian Intermediate EFL Learners' Motivation toward Learning English Vocabularies. *Procedia-Social and Behavioral Sciences*, *98*, 1738-1743.

Shen, H. H. (2010). Imagery and verbal coding approaches in Chinese vocabulary instruction. *Language Teaching Research, 14*(4), 485-499.

Singleton, D. (1999). *Exploring the second language mental lexicon*. Cambridge: Cambridge University Press.

Sökmen, A. J. (1997). Current trends in teaching second language vocabulary. in N. Schmitt & M. McCarthy (eds.), *Vocabulary: description, acquisition and pedagogy* (pp. 237-257). Cambridge: Cambridge U.P.

Someren, M. V., Barnard, Y. F., & Sandberg, J. A. (1994). *The think aloud method: a practical approach to modelling cognitive processes*. London: Academic Press.

Squire, K. (2003). Video games in education. *Int. J. Intell. Games & Simulation*, *2*(1), 49-62.

Squire, K. (2005). Changing the game: What happens when video games enter the classroom?. *Innovate: Journal of online education*, *1*(6), 5.

Squire, K. D. (2008). Video game–based learning: An emerging paradigm for instruction. *Performance Improvement Quarterly*, *21*(2), 7-36.

Suh, S., Kim, S. W., & Kim, N. J. (2010). Effectiveness of MMORPG-based instruction in elementary English education in Korea. *Journal of Computer Assisted Learning*, *26*(5), 370-378.

Sundqvist, P. (2019). Commercial-off-the-shelf games in the digital wild and L2 learner vocabulary. *Journal of Language Learning & Technology, 23*(1), 87-113.

Sundqvist, P., & Sylvén, L. K. (2012). World of VocCraft: Computer games and Swedish learners L2 English vocabulary. *Digital Games in Language Learning and Teaching (New Language Learning and Teaching Environments)*, 189-208.

Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. Gass & C.

Madden (Eds.), *Input in second language acquisition* (pp. 235–253). Rowley, MA: Newbury House.

Sweller, J. (2010). Cognitive Load Theory: Recent Theoretical Advances. In: Jan L. Plass et al. (eds.) *Cognitive Load Theory*. pp. 29-47. [Online]. Cambridge: Cambridge University Press. Available from: Cambridge.

Taevs, M., Dahmani, L., Zatorre, R. J., & Bohbot, V. D. (2010). Semantic elaboration in auditory and visual spatial memory. *Frontiers in psychology*, *1*, 228.

Tahmasbi, M., & Farvardin, M. T. (2017). Probing the Effects of Task Types on EFL Learners' Receptive and Productive Vocabulary Knowledge: The Case of Involvement Load Hypothesis. *SAGE Open*, *7*(3), 2158244017730596.

Thuy, N. (2013). The effects of semantic mapping on vocabulary memorizing. *Foreign Language Learning and Teaching*, 2, 628- 659. Retrieved from www.litu.tu.ac.th/journal/FLLTCP/Proceeding/628.pdf.

Truscott, J. (1998). Noticing in second language acquisition: A critical review. *second language Research*, *14*(2), 103-135.

Tseng, W. T., & Schmitt, N. (2008). Toward a model of motivated vocabulary learning: A structural equation modeling approach. *Language Learning*, *58*(2), 357-400.

Tsubaki, M. (2007). The involvement load hypothesis: An inquiry into vocabulary learning. *Journal of Bunkyo Gakuin University*, (6), 175-184.

Vahdat, S., & Rasti Behbahani, A. R. (2013). The effect of video games on Iranian EFL learners' vocabulary learning. *Reading*, *13*(1), 61-71.

Vaismoradi, M., Jones, J., Turunen, H., & Snelgrove, S. (2016). Theme development in qualitative content analysis and thematic analysis. *Journal of Nursing Education and Practice*, *6*(5), 100.

Vaismoradi, M., Turunen, H., & Bondas, T. (2013). Content analysis and thematic analysis: Implications for conducting a qualitative descriptive study. *Nursing & health sciences*, *15*(3), 398-405.

Webb, S. (2007a). Learning word pairs and glossed sentences: The effects of a single context on vocabulary knowledge. *Language teaching research*, *11*(1), 63-81.

Webb, S. (2007b). The effect of repetition on vocabulary knowledge. *Applied Linguistics, 28*(1), 46-65.

Webb, S. (2008). The effects of context on incidental vocabulary learning. *Reading in a foreign Language*, *20*(2), 232-245.

Webb, S., & Chang, A. C. (2015). Second language vocabulary learning through extensive reading with audio support: How do frequency and distribution of occurrence affect learning? *Language Teaching Research*, *19*(6), 667-686.

Wessels, S. (2011). Promoting vocabulary learning for English learners. *The Reading Teacher*, *65*(1), 46-50.

Wu, S., Franken, M., & Witten, I. H. (2012). Collocation games from a language corpus. In *Digital games in language learning and teaching* (pp. 209-229). Palgrave Macmillan, London.

Xie, H., Zou, D., Wang, F. L., & Wong, T. L. (2017, June). A Review on Recent Development of the Involvement Load Hypothesis. Paper presented at *International Conference on Blended Learning*, 447-452

Xu, X. (2010). The effects of glosses on incidental vocabulary acquisition in reading. *Journal of language teaching and research*, *1*(2), 117-120.

Yaqubi, B., Rayati, R. A., & Allemzade Gorgi, N. (2012). The involvement load hypothesis and vocabulary learning: The effect of task types and involvement index on L2 vocabulary acquisition. *Journal of Teaching Language Skills*, *29*(1), 145-163.

Yip, F. W., & Kwan, A. C. (2006). Online vocabulary games as a tool for teaching and learning English vocabulary. *Educational media international*, *43*(3), 233-249.

Yoshida, M. (2008). Think-aloud protocols and type of reading task: The issue of reactivity in L2 reading research. In *Selected proceedings of the 2007 second language research forum* (Vol. 199209). Somerville, MA: Cascadilla Proceedings Project.

Yoshii, M. (2006). L1 and L2 glosses: Their effects on incidental vocabulary learning. *Language Learning & Technology, 10*(3), 85-101.

Yule, G. (2014). *The study of language (5th Ed.)*. Cambridge: Cambridge University Press.

Zhonggen, Y. (2018). Differences in serious game-aided and traditional English vocabulary acquisition. *Computers & Education*, *127*, 214-232.

Zimmerman, C. B. (1994). *Self-selected reading and interactive vocabulary instruction: Knowledge and perceptions of word learning among L2 learners* (Unpublished doctoral dissertation, University of Southern California, Los Angeles).

Zin, N. A. M., Jaafar, A., & Yue, W. S. (2009). Digital game-based learning (DGBL) model and development methodology for teaching history. *WSEAS transactions on computers*, *8*(2), 322-333.

Zou, D. (2016). Comparing dictionary-induced vocabulary learning and inferencing in the context of reading. *Lexikos*, *26*(1), 372-390.

Zou, D. (2017). Vocabulary acquisition through cloze exercises, sentence-writing and composition-writing: Extending the evaluation component of the involvement load hypothesis. *Language teaching research*, *21*(1), 54-75.

Zou, D., & Xie, H. (2018). Personalized word-learning based on technique feature analysis and learning analytics. *Journal of Educational Technology & Society*, *21*(2), 233-244.

APPENDIX A: VOCABULARY SIZE TEST (PERSIAN)

First 1000

1. SEE: They **saw** it.
a.بریدن
b.منتظر ماندن
c.نگاه کرد
d.شروع کرد

2. TIME: They have a lot of **time**.
a.پول
b.غذا
c.ساعت ها
d.دوستان

3. PERIOD: It was a difficult **period**.
a. سوال
b. زمان
c. وظایف
d. کتاب

4. FIGURE: Is this the right **figure**?
a. جواب
b. مکان
c. زمان
d. عدد

5. POOR: We are **poor**.
a. بی پول
b. خوشحال بودن
c. مشتاق بودن
d. از کار سخت راضی نبودن

6. DRIVE: He **drives** fast.
a. شنا کردن
b. یادگرفتن
c. توپ پرتاب کردن
d. از ماشین استفاده کردن

7. JUMP: She tried to **jump**.
a. روی آب دراز کشیدن
b. از زمین سریع برخاستن
c. نگه داشتن ماشین بر لبه جاده
d. سریع حرکت کردن

8. SHOE: Where is your **shoe**?
a. کسی که مراقب شماست
b. چیزی که پولت رو توش نگه داری میکنی
c. چیزی که باهاش مینویسی
d. چیزی که میپوشی پاهات

9. STANDARD: Her **standards** are very high.
a. چیزای کوچیکی که کف کفش هستند
b. نمره هایی که تو مدرسه گرفته
c. پولی که درخواست کرده
d. حدی که هرکسی در هرچیزی بهش میرسه

10. BASIS: This was used as the **basis**.
a. جواب

Second 1000

11. MAINTAIN: Can they **maintain** it?
a. ثابت نگه داشتن
b. بزرگ تر کردن
c. یکی بهتر از قبلی برداشتن
d. گرفتن

12. STONE: They sat on a **stone**.
a. یک چیز سخت
b. یک نوع صندلی
c. چیز نرمی روی کف خانه
d. قسمتی از یک درخت

13. UPSET: I am **upset**.
a. خسته
b. مشهور
c. پولدار
d. ناراحت

14. DRAWER: The **drawer** was empty.
a. جعبه ای که سر میخوره
b. جایی که ماشین ها نگهداری میشود
c. جایی برای سرد نگه داشتن
d. کلبه حیوانات

15. PATIENCE: He has no **patience**.
a. منتظر ماندن
b. وقت آزاد
c. ایمان
d. نداستن درست و غلط

16. NIL: His mark for that question was **nil**.
a. خیلی بد
b. هیچ چیز
c. خیلی خوب
d. در مرکز

17. PUB: They went to the **pub**.
a. جایی که مردم مینوشند و صحبت میکنند
b. جایی که از پول نگهداری میکند
c. جایی برای سرد نگه داشتن
d. لانه حیوانات

18. CIRCLE: Make a **circle**.
a. شکل بی زانویه
b. جای خالی
c. شکل گرد
d. سوراخ بزرگ

19. MICROPHONE: Please use the **microphone**.
a. وسیله ای برای گرم کردن غذا
b. دستگاهی برای بلندتر کردن صدا
c. دستگاهی که در آن چیزها بزرگتر به نظر میرسند
d. تلفن همراه

20. PRO: He's a **pro**
a. کسی که برای یافتن معماها استخدام شده

b. جایی برای استراحت
c. قدم بعدی
d. قسمت اصلی

Third 1000

21. SOLDIER: He is a **soldier**.
a. تاجر
b. دانش آموز
c. کسی که از ابزار آهنی استفاده می کند
d. کسی که در ارتش خدمت میکند

22. RESTORE: It has been **restored**.
a. دوباره گفته شده
b. به فرد دیگری سپرده شده
c. قیمت کمتری گرفته
d. دوبار نو کردن

23. JUG: He was holding a **jug**.
a. جایی برای ریختن مایعات
b. بحث غیر جدی
c. کلاه لبه دار نرم
d. وسیله جنگی که منفجر می شود

24. SCRUB: He is **scrubbing** it.
a. کندن خط های نازک در چیزی
b. تعمیر کردن
c. ساویدن چیزی برای تمیزتر شدن
d. طرح ساده ای از چیزی کشیدن

25. DINOSAUR: t-Rex is a **dinosaur**
a. دزدان دریایی
b. موجودات کوچکی با بدنی مانند انسان و بال
c. موجودات بسیار بزرگ که آتش از دهانشان خارج می شود
d. موجوداتی که سال ها پیش میزیسته اند

26. STRAP: He broke the **strap**.
a. وعده دادن
b. درپوش بالایی
c. بشقاب کم غذا
d. بندی برای بستن دور و نگهداشتن اشیا به همدیگر

27. PAVE: It was **paved**.
a. جلوگیری از پیشرفتن
b. تقسیم شده
c. دور چیزی را طلایی کردن
d. با سطح سختی پوشانده شدن

28. DASH: They **dashed** over it.
a. سریع حرکت کردن
b. آرام حرکت کردن
c. جنگیدن
d. سریع نگاه کردن

29. ROVE: He couldn't stop **roving**.
a. مست شدن
b. سفر کردن
c. با دهان آهنگ زدن
d. سخت کار کردن

b. فرد نادان
c. کسی که برای روزنامه مطلب مینویسد
d. کسی که برای بازی در ورزشی پول میگیرد

Fourth 1000

31. COMPOUND: They made a new **compound**.
a. توافق کردن
b. چیزی که از دو یا سه قسمت تهیه شده
c. گروه تجاری
d. براساس تجارب گذشته چیزی را حدس زدن

32. LATTER: I agree with the **latter**.
a. فردی از اعضای کلیسا
b. دلایل ارایه شده
c. آخرین
d. جواب

33. CANDID: Please be **candid**.
a. مراقب بودن
b. همدلی کردن
c. عادل بودن
d. گفتن هرچه در ذهن داشتن

34. TUMMY: Look at my **tummy**.
a. چیزی برای پوشش سر
b. شکم
c. جانور کوچک پشمالو
d. انگشت

35. QUIZ: We made a **quiz**.
a. چیزی برای نگه داشتن تیر
b. اشتباه بزرگ
c. یک سری سوال
d. جعبه ای برای لانه پرندگان

36. INPUT: We need more **input**.
a. وارد کردن اطلاعات ، برق ، یا هرچیزی در چیز دیگر
b. کارگران
c. چیزی برای پرکردن سوراخ های چوب
d. پول

37. CRAB: Do you like **crabs**?
a. موجود دریایی که از بغل حرکت می کند
b. کیک های بسیار نازک و کوچک
c. یقه کوچک و سفت
d. حشره کوچکی که شب ها صدا می دهد

38. VOCABULARY: You'll need more **vocabulary**.
a. کلمات
b. مهارت
c. پول
d. اسلحه

39. REMEDY: We found a good **remedy**.
a. راهی برای حل مشکل
b. جایی برای غذا خوردن در عموم
c. راهی برای آماده سازی غذا
d. قوانین اعداد

30. LONESOME: He felt **lonesome**.
a. ناشکری کردن
b. بسیار خسته
c. تنها و بی کس
d. پر از انرژی

Fifth 1000

41. DEFICIT: The company had a large **deficit**.
a. بیش از درآمد خرج کردن
b. از ارزش چیزی بسیار کم شدن
c. برنامه ای برای خرج کردن داشتن
d. پول زیادی در بانک داشتن

42. WEEP: He **wept**.
a. تمام کردن درس
b. گریه کردن
c. مُردن
d. نگران بودن

43. NUN: We saw a **nun**
a. موجود دراز و لاغری که در زمین زندگی میکند
b. تصادف وحشتناک
c. زنی که بسیار مذهبی است
d. نور روشن ناشناخته در آسمان

44. HAUNT: The house is **haunted**.
a. پر از تزیینات
b. کرایه کردن
c. خالی کردن
d. پر از ارواح

45. COMPOST: We need some **compost**.
a. پشتیبانی شدید
b. به کسی کمک کردن تا حالش بهتر شود
c. چیز سختی که از سنگ و شن چسبیده به همدیگر درست شده
d. ماده گیاهی گندیده

46. CUBE: I need one more **cube**.
a. چیز تیزی که برای دوختن استفاده می شود
b. مکعب سخت و جامد
c. استکان بلند بدون نعلبکی
d. تکه کاغذ از وسط تا شده

47. MINIATURE: It is a **miniature**.
a. یک چیز نازک و ظریف
b. وسیله ای برای دیدن اشیاء ریز
c. موجودات ریز زنده
d. خط کوچکی که کلمات را در نوشتن به هم متصل میکند

48. PEEL: Shall I **peel** it?

a. برای مدتی در آب گذاشتن
b. پوست چیزی را کندن
c. سفید کردن
d. خلال کردن

49. FRACTURE: They found a **fracture**.
a. تیکه شکسته
b. تیکه کوچک

40. ALLEGE: They **alleged** it.
a. بدون مدرک ثابت کردن
b. دزدیدن ایده کسی برای دیگری
c. بیان حقایقی برای اثبات ادعا
d. بحث کردن بر حقایقی که ایده را پشتیبانی میکنند

Sixth 1000

51. DEVIOUS: Your plans are **devious**.
a. حقه باز
b. خوش ساخت
c. سرسری گرفته شده
d. بسیار گرانتر از حد معمول

52. PREMIER: The **premier** spoke for an hour.
a. کسی که در سطح پایین دادگاه کار میکند
b. استاد دانشگاه
c. ماجراجو
d. رییس دولت

53. BUTLER: They have a **butler**.
a. خدمتکار مرد
b. دستگاه بریدن درختان
c. معلم خصوصی
d. اتاق تاریک و سردی زیر خانه

54. ACCESSORY: They gave us some **accessories**.
a. برگه ای که با آن می توان وارد کشوری شد
b. دستورات رسمی
c. ایده های منتخب
d. قطعات اضاف و مرتبط

55. THRESHOLD: They raised the **threshold**.
a. پرچم
b. نقطه یا خطی که در آنجا چیزی عوض میشود
c. سقفی درون ساختمان
d. هزینه قرض گرفتن پول

56. THESIS: She has completed her **thesis**.
a. پایان نامه
b. آخرین سخنرانی قاضی در پایان مراسم محاکمه
c. اولین سال کار معلمی
d. دروهای اضافه آموزش درمانی و پرستاری

57. STRANGLE: He **strangled** her.
a. با فشار دادن گلو کسی را کشتن
b. تمام آنچه که خواسته شده را دادن
c. به زور او را دور کردن
d. بسیار کسی را تحسین کردن

58. CAVALIER: He treated her in a **cavalier** manner.
a. بی احتیاط
b. مودبانه
c. عجب و غریبانه
d. مانند یک برادر رفتار کردن

59. MALIGN: His **malign** influence is still felt.
a. شیطان
b. خوب

c. کُت کوچک
d. جواهر کم یاب

50. BACTERIUM: They find a single **bacterium** it.
a. موجودات کوچکی که باعث مریضی میشوند
b. گیاهی با گل های سرخ و نارنجی
c. موجودی که در پشتش آب حمل میکند
d. چیزی که دزدیده شده و به مغازه ای فروخته شده

Seventh 1000

61. OLIVE: We bought **olives**.
a. میوه روغنی
b. گلو صورتی یا سرخ خوشبو
c. لباس شنا برای مردان
d. ابزاری برای برون کشیدن گیاهان هرزه

62. QUILT: They made a **quilt**.
a. وصیت نامه
b. توافق قطعی
c. روپوش ضخیم و گرم برای تخت
d. قلم پری

63. STEALTH: We did it by **stealth**.
a. پول زیادی خرج کردن
b. به زور از کسی اعتراف گرفتن
c. مخفیانه و دزدکی حرکت کردن
d. از مشکلات مواجه شده درس نگرفتن

64. SHUDDER: The boy **shuddered**.

a. با صدای کم صحبت کردن
b. تقریبا پُر
c. لرزیدن
d. به بلندی فراخوانده شدن

65. BRISTLE: The **bristles** are too hard.
a. سوالات
b. موهای گندمی کوتاه
c. تخت تاشو
d. کف کفش

66. BLOC: They have joined this **bloc**.
a. گروه موسیقی
b. گروه دزدان
c. گروه کوچک سربازی که جلوتر از بقیه فرستاده شده اند
d. گروه کشورهای مشترک در یک هدف

67. DEMOGRAPHY: This book is about **demography**.
a. مطاله زمین و تغییرات مرتبط با آن
b. مطالعات عکس شناسی و درک اعداد
c. مطالعات آب
d. مطالعات اجتماعی

68. GIMMICK: That's a good **gimmick**.
a. چیزی که هنگام کار در بلندی بر روی آن می ایستند
b. چیز کوچکی با جایی برای نگهداری پول
c. کاری برای جلب توجه دیگران

c. بسیار مهم
d. راز

60. VEER: The car **veered**.

a. ناگهان به جهت دیگر رفتن
b. به لرزه درآمدن
c. صدای بسیار بلندی ایجاد کردن
d. بدون چرخیدن تایرها از بغل سر خوردن

Eighth 1000

71. ERRATIC: He was **erratic**.
a. بدون خطا
b. خیلی بد
c. بسیار مودب
d. بی ثبات و بدون آمادگی

72. PALETTE: He lost his **palette**.
a. سبدی برا حمل ماهی
b. میل خوردن غذا
c. زن همراه و جوان
d. تخته نقاش برای مخلوط کردن رنگ ها

73. NULL: His influence was **null**.
a. نتیجه خوبی داشتن
b. بدردبخور نبودن
c. تاثیری نداشتن
d. اثر طولانی داشتن

74. KINDERGARTEN: This is a good **kindergarten**.
a. فعالیتی که در فراموش کردن نگرانی ها کمک می کند
b. مکانی برای یادگیری
c. کیف بزرگ و محکمی که روی پشت حمل می شود
d. جایی که از آن کتاب قرض می کنید

75. ECLIPSE: There was an **eclipse**.
a. باد قوی
b. صدای بلند برخورد چیزی با آب
c. کشتن تعداد زیادی از انسان ها
d. وقتی خورشید پشت ماه پنهان می شود

76. MARROW: This is the **marrow**.
a. سمبل خوش شانسی
b. قسمت نرم وسط استخوان
c. کنترل هواپیما
d. افزایش حقوق

77. LOCUST: There were hundreds of **locusts**.

a. حشره بالدار
b. کمک های غیرمالی
c. افرادی که گوشت نمی خورند
d. گل های وحشی با رنگ براق

78. AUTHENTIC: It is **authentic**.
a. واقعی
b. بسیار پر سر و صدا
c. قدیمی

d. حُقه هوشیارانه

69. AZALEA: This **azalea** is very pretty.
a. گیاه کوچکی با گل های دسته ای
b. مواد سبکی که از مواد طبیعی ساخته شده
c. لباس هایی که زنان هندی به تن می کنند
d. صدف های دریایی که مانند بادبزن هستند

70. YOGHURT: This **yoghurt** is disgusting.
a. گل سیاهی که در ته رودخانه ها پیدا میشوند
b. ناسالم و غیربهداشتی
c. ماده سفت شده و غلیظی که از شیر گرفته میشود
d. میوه بنفش بزرگ با گوشت نرم

Ninth 1000

81. HALLMARK: Does it have a **hallmark**?
a. تمبَری که نشان میدهد کی استفاده شود
b. تمبَری که نشان کیفیت است
c. نشان تایید سلطنتی
d. علامت یا نشان کپی ممنوع

82. PURITAN: He is a **puritan**.
a. کسی که توجه دوست دارد
b. فردی بسیار پایبند به اخلاقیات
c. فرد خانه به دوش
d. آدم خسیس

83. MONOLOGUE: Now he has a **monologue**.

a. عینک تک چشم
b. صحبت کردم پیوسته و بدون بریدگی
c. بالاترین منسب قدرت
d. عکسی که با چسباندن حروف به همدیگر بوجود آمده

84. WEIR: We looked at the **weir**.
a. کسی که رفتار عجیبی دارد
b. جای گلی و تر که گل های آبی در آن است
c. ساز موسیقایی بادی ساخته شده از فلز
d. چیزی که روی رودخانه کشیده می شود تا آب پشتش جمع شود

85. WHIM: He had lots of **whims**.
a. سکه های طلای قدیمی
b. اسب ماده
c. ایده های عجیب و بی انگیزه
d. برجستگی های قرمز دردناک

86. PERTURB: I was **perturbed**.
a. به توافق دست یافتن
b. نگران
c. بسیار گیج کننده
d. بسیار خیس

87. REGENT: They chose a **regent**.
a. فرد بی مسولیت
b. مُجری
c. قانون گذاری که به جای پادشاه منسوب شده است
d. نماینده

88. OCTOPUS: They saw an **octopus**.
a. پرنده بزرگی که شب ها شکار می کند

d. مانند صحرا

79. CABARET: We saw the **cabaret**.
a. سوراخ توی دیوار را با رنگ پوشاندن
b. محل ساز و رقص
c. حشره کوچک خزنده
d. موجودی که نیمه ماهی و نیمه انسان است

80. MUMBLE: He started to **mumble**.
a. عمیق تفکر کردن
b. دیوانه وار تکان دادن
c. پشت کسی قایم شدن
d. ناواضح صحبت کردن

Tenth 1000

91. AWE: They looked at the mountain with **awe**.
a. نگرانی
b. علاقه
c. تعجب و حیرت
d. احترام

92. PEASANTRY: He did a lot for the **peasantry**.
a. افراد بومی
b. مکان عبادت
c. پاتوق بازرگانان
d. کشاورزان فقیر

93. EGALITARIAN: This organization is **egalitarian**.
a. اطلاعات زیادی از خود پخش نکردن
b. تغییرات ناپسند
c. درخواست قضاوت فرستادن به دادگاه
d. با همه کارگران به مساوات رفتار کردن

94. MYSTIQUE: He has lost his **mystique**.
a. تن سالم
b. دیگران را گول زدن و نشان دادن که قدرت فراطبیعی دارند
c. معشوقه یک مرد زن دار بودن
d. موهای پشت لب

95. UPBEAT: I'm feeling really **upbeat** about it.
a. ناراحت
b. خوب
c. آزار دادن
d. گیج بودن

96. CRANNY: We found it in the **cranny**!
a. فروش اشیا ناخواسته
b. گذرگاه باریک
c. انبار زیر سقف یا زیر خانه
d. جعبه چوبی بزرگی

97. PIGTAIL: Does she have a **pigtail**?
a. دسته موی بافته شده
b. لباس های زیادی که پشت یک لباس زنانه آویزان شده
c. گیاهی با گلهای صورتی سفید که بر روی بوته روبه پایین هستند
d. معشوق

98. CROWBAR: He used a **crowbar**.
a. تیکه آهن دراز با ته خم شده

b. زیر دریایی
c. ماشینی که با چرخاندن تیغه ها پرواز می کند
d. موجود دریایی یا هشت پا

89. FEN: The story is set in the **fens**.
a. زمین پستی که با آب پوشیده شده
b. قسمتی از زمین بلند با درختان کم
c. قسمتی از شهر با خانه های بی کیفیت و فقیر
d. زمان های بسیار کهن

90. LINTEL: He painted the **lintel**.
a. تیرآهن بالای در یا پنجره
b. کشتی کوچکی برای رسیدن به خشکی از کشتی بزرگتر
c. درخت زیبایی با شاخه های زیاد و میوه های سبز
d. تخته ای که نشان دهنده صحنه در تئاتر است

Eleventh 1000

101. EXCRETE: This was **excreted** recently.
a. فشار داده شده یا به بیرون فرستاده شده
b. واضح کردن
c. اکتشاف علمی
d. لیست کردن اجناس غیرقانونی

102. MUSSEL: They bought **mussels**.
a. توپ شیشه ای کوچک برای بازی
b. ماهی صدف دار
c. میوه های بزرگ بنفش
d. تیکه کاغذی برای تمیز نگهداشتن لباس ها هنگام غذا خوردن

103. YOGA: She has started **yoga**.
a. کاردستی با گره
b. نوعی ورزش برای بدن وذهن
c. بازی بتمینتون
d. نوعی رقص از کشورهای شرقی

104. COUNTERCLAIM: They made a **counterclaim**.
a. ادعای فردی در یک پرونده قضایی که با فرد دیگری برابر است
b. درخواست پس گرفتن جنس خراب
c. قرارداد بین دو کمپانی برای معاوضه کردن
d. روتختی

105. PUMA: They saw a **puma**.
a. خانه کوچک ساخته شده از کلوخ
b. درخت سرزمین های خشک و گرم
c. بادی قوی که همه چیز در راه خود می بلعد
d. گربه بزرگ وحشی

106. PALLOR: His **pallor** caused them concern.
a. دمای بالای بدن و نامعمول
b. بی علاقگی به همه چیز
c. گروه دوستان
d. بی رنگی پوست

107. APERITIF: She had an **aperitif**.
a. صندلی تک دسته ای
b. معلم آواز خصوصی
c. کلاه بزرگ با پر بلند

b. نام جعلی
c. وسیله تیزی برای ایجاد سوراخ در چرم
d. عصای فلزی سبک

99. RUCK: He got hurt in the **ruck**.
a. گودی بین شکم و بالای پا
b. هل دادن و فشار آوردن
c. گروه بازیکنانی که دور یک توپ جمع شده اند
d. در یک زمین برفی مسابقه دادن

100. LECTERN: He stood at the **lectern**.
a. میزی برای نگه داشتن کتاب برای مطالعه
b. میز یا تخته ای که برای قربانی های کلیسا استفاده می شود
c. جایی که از آن نوشیدنی خریداری می شود
d. روی لبه

Twelfth 1000

111. HAZE: We looked through the **haze**.
a. پنجره گرد کشتی
b. هوای ناصاف
c. بریده های چوب یا پلاستیک برای پوشاندن پنجره
d. لیست اسامی

112. SPLEEN: His **spleen** was damaged.
a. استخوان زانو
b. ارگانی که نزدیک شکم است
c. لوله فاضلاب
d. احترام به خود

113. SOLILOQUY: That was an excellent **soliloquy**.
a. آهنگی برای شش نفر
b. جمله حکیمانه
c. ایجاد سرگرمی با نور و موسیقی
d. جملات فرد تنها در صحنه تئاتر

114. REPTILE: She looked at the **reptile**.
a. کتاب خطی قدیمی
b. موجود خونسرد و سخت پوست
c. فروشنده دوره گرد
d. عکسی که با چسباندن تیکه های کاغذ رنگی درست می شود

115. ALUM: This contains **alum**.
a. ماده سمی از یک گیاه همه جا روی
b. ماده نرمی که از الیاف مصنوعی درست شده
c. پودر تنباکو که در بینی فرو می کنند
d. ترکیبات شیمیایی که معمولا شامل آلمینیوم هستند

116. REFECTORY: We met in the **refectory**.
a. اتاق غذاخوری
b. اداره ای برای امضا اوراق رسمی
c. خوابگاه چند تخته
d. گلخانه شیشه ای

117. CAFFEINE: This contains a lot of **caffeine**.
a. ماده خواب آور
b. الیاف بدست آمده از برگ های سخت
c. ایده های اشتباه

d. نوشیدنی قبل از غذا

108. HUTCH: Please clean the **hutch**.
a. توری برای تمیز نگه داشتن آب
b. جایی پشت ماشین برای اساسیه
c. مکان آهنین در وسط چرخ های دوچرخه
d. قفسی برای موجودات کوچک

109. EMIR: We saw the **emir**.
a. پرنده ای با پرهای خمیده بلند
b. زنی که در کشورهای شرقی از کودکان دیگران مراقبت می کند
c. فرد دارای قدرت و ریاست در کشورهای خاورمیانه
d. خانه ای که از تکه های یخ ساخته شده

110. HESSIAN: She bought some **hessian**.
a. ماهی صورتی روغنی
b. چیزهایی که باعث ایجاد حس خوش در ذهن می شوند
c. لباس زبر
d. ریشه ای با طعم بسیار قوی برای مزه دار کردن غذاها

Thirteenth 1000

121. UBIQUITOUS: Many weeds are **ubiquitous**.
a. براحتی از بین برده نمی شود
b. دارای ریشه بلند و طولانی
c. در بیشتر کشورها پیدا می شود
d. در زمستان نمی روید

122. TALON: Just look at those **talons**!
a. بلندی کوهستان
b. ناخن خمیده و تیز پرندگان شکاری
c. پوشش زرهی محکم
d. افرادی که خودشان را مضحکه میکنند بدون اینکه بدانند

123. ROUBLE: He had a lot of **roubles**.
a. سنگ قرمز بسیار با ارزش
b. افراد دور خانواده
c. پول روسی
d. مشکلات اخلاقی یا روانی در ذهن

124. JOVIAL: He was very **jovial**.
a. از نظر اجتماعی سطح پایین
b. منتقد دیگران
c. پر از شادی و خوشی
d. دوستانه

125. COMMUNIQUE: I saw their **communique**.
a. گزارش منتقدانه در مورد یک موسسه
b. باغی که متعلق به افراد زیادی از یک جماعت است
c. کاغذهای تبلیغاتی
d. اعلان رسمی

126. PLANKTON: We saw a lot of **plankton**.
a. گیاهان سمی که به سرعت گسترش می یابند
b. گیاهان و حیوانات کوچکی که در آب هستند
c. درختانی که برای ساختن چوب های سخت استفاده می شوند

d. ماده ای انرژی زا

118. IMPALE: He nearly got **impaled**.
a. با خشونت جریمه شدن
b. در زندان انداختن
c. با وسیله تیزی سوراخ کردن
d. درگیر یک دعوا بودن

119. COVEN: She is the leader of a **coven**.
a. گروه خواننده کوچک
b. تجارتی که کارگران همان شرکت آن را خریداری می کنند
c. جامعه مخفی
d. گروهی از زنان کلیسا که بسیار مذهبی هستند

120. TRILL: He practiced the **trill**.
a. قطعه تزینی از یک قطعه موسیقی
b. نوعی وسیله نخی
c. شیوه پرتاب توپ
d. نوعی رقص و چرخش بر روی انگشت

Fourteenth 1000

131. CANONICAL: These are **canonical** examples.
a. نمونه هایی که در آن ها قوانین رایج شکسته شده است
b. مثال هایی که از یک کتاب مقدس استخراج شده است
c. مثال های معمولی و رایج
d. نمونه هایی که اخیرا کشف شده اند

132. ATOP: He was **atop** the hill.
a. در ته چیزی
b. در بالای چیزی
c. در کنار چیزی
d. آن طرف تر از چیزی بودن

133. MARSUPIAL: It is a **marsupial**.
a. موجودی با پاهای سخت
b. گیاهی که رشدش متوقف نمی شود
c. گیاهی که گلش به سمت خورشید می چرخد
d. موجودی با کیسه برای حمل بچه

134. AUGUR: It **augured** well.
a. چیزهای خوب وعده داده شده برای آینده
b. با چیزی که در انتظار داشتن موافق بودن
c. رنگی که با رنگ دیگری جور درمی آید
d. آهنگی با صدای زیبا و رسا

135. BAWDY: It was very **bawdy**.
a. غیرقابل پیش بینی
b. لذت بخش
c. تهاجمی
d. گستاخ

136. GAUCHE: He was **gauche**.
a. پرحرف
b. انعطاف پذیر
c. عجیب و غریب

d. گِل خاکستری که باعث ریزش یا لغزندگی زمین می شوند

127. SKYLARK: We watched a **skylark**.
a. نمایش هواپیماها
b. وسیله ساخته شده دست بشر که دور زمین می چرخد
c. فردی که کارهای طنزآمیز انجام می دهد
d. پرنده ای که که در حین پرواز در بلندی آواز می خواند

128. BEAGLE: He owns two **beagles**.
a. ماشین های تندرو که سقف تاشو دارند
b. تفنگ بزرگی که می تواند تیرهای زیادی همزمان شلیک کند
c. سگ کوچک با گوش های دراز
d. خانه هایی که در محل های جشن قرار گرفته اند

129. ATOLL: The **atoll** was beautiful.
a. جزیره پست مرجانی که در اطراف دریاچه منتهی به دریا ساخته شده
b. اثر هنری بافته شده از الیاف باکیفیت با رنگ ها و طرح های مختلف
c. تاج کوچک پر از جواهری که زن ها عصر بر سر می گذارند
d. جای باریک و پر از سنگی که رودخانه از آن عبور می کند

130. DIDACTIC: The story is very **didactic**.
a. سخت تلاش کردن تا چیزی را به کسی درس دادن
b. بسیار غیرقابل باور
c. مرتبط به اعمال مهیج
d. نامفهوم و مبهم نوشته شده

d. قاطع

137. THESAURUS: She used a **thesaurus**.
a. نوعی دیکشنری
b. ترکیب شیمیایی
c. روش خاص صحبت کردن
d. تزریق زیر پوست

138. ERYTHROCYTE: It is an **erythrocyte**.
a. داروی کاهش دهنده درد
b. قسمت قرمز خون
c. فلز سفید – قرمز
d. عضوی از گونه وال ها

139. CORDILLERA: They were stopped by the **cordillera**.
a. قانونی خاص
b. کشتی مسلح
c. رشته کوه
d. پسر بزرگتر پادشاه

140. LIMPID: He looked into her **limpid** eyes.
a. شفاف
b. پر از اشک
c. قهوه ای پررنگ
d. زیبا

APPENDIX B: RECEPTIVE TEST

1. Match the words with their meanings by numbers:

A:

1. Debris
2. Shack _____ چفت در
3. Latch _____ خُرد و ریز
4. Hook _____ قلاب
5. Skull
6. Matches

B:

1. Fragments
2. Portrait
3. Ember _____ زغال داغ
4. Sign _____ تیکه خُرده
5. Magnifier _____ فَرش
6. Carpet

2. Choose the best meaning for the given words

Portrait means:

a) قاب عکس b) فَرش c) کبریت d) جمجمه

Drape means:

a) حصار فلزی b) ذره بین c) پرده توری d) کشو

Magnifier means:

a) ذره بین b) کشو c) زغال داغ d) دنده ها

Closet means:

a) سرپوش b) قاب عکس c) زنجیر d) گنجه

3. Translate the following words into Persian

Skull _____
Matches _____
Cover _____
Sign _____
Fence-barriers _____
Chain _____
Gears _____
Drawer _____
Shack _____
Burner _____

APPENDIX C: PRODUCTIVE TEST

1. Match the words with their meanings by numbers:

A:

1. قلاب
2. فَرش
3. چفت در _____ Fragments
4. گنجه _____ Latch
5. خُرد و ریز _____ Debris
6. تیکه خُرده

B:

1. ذره بین
2. فَرش
3. زنجیر _____ Carpet
4. حصار فلزی _____ Drape
5. پرده توری _____ Magnifier
6. کشو

C:

1. دنده ها
2. نشانه _____ Chain
3. گنجه _____ Sign
4. پوشش _____ Gears
5. کشو _____ Cover
6. زنجیر

2. Fill in the blanks with appropriate words.

- They were poor; they were living in a small sh_____.
- I got a big fish with a long ho_____.
- A por_____ of my grand-father is on the wall of living room.
- In her cl_____, Alice puts her money and golds.

3. Translate the following words into English.

جمجمه _____

کبریت _____

گاز پیک نیک _____

کشو _____

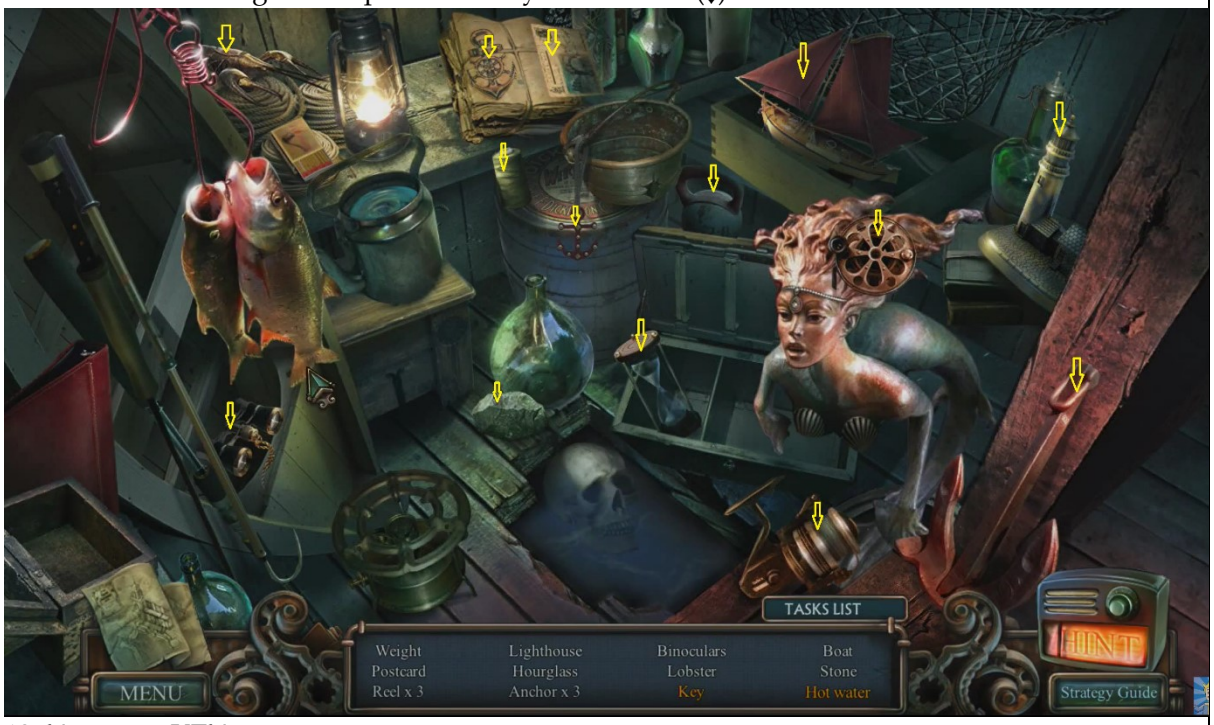
زغال داغ _____

حصار فلزی _____

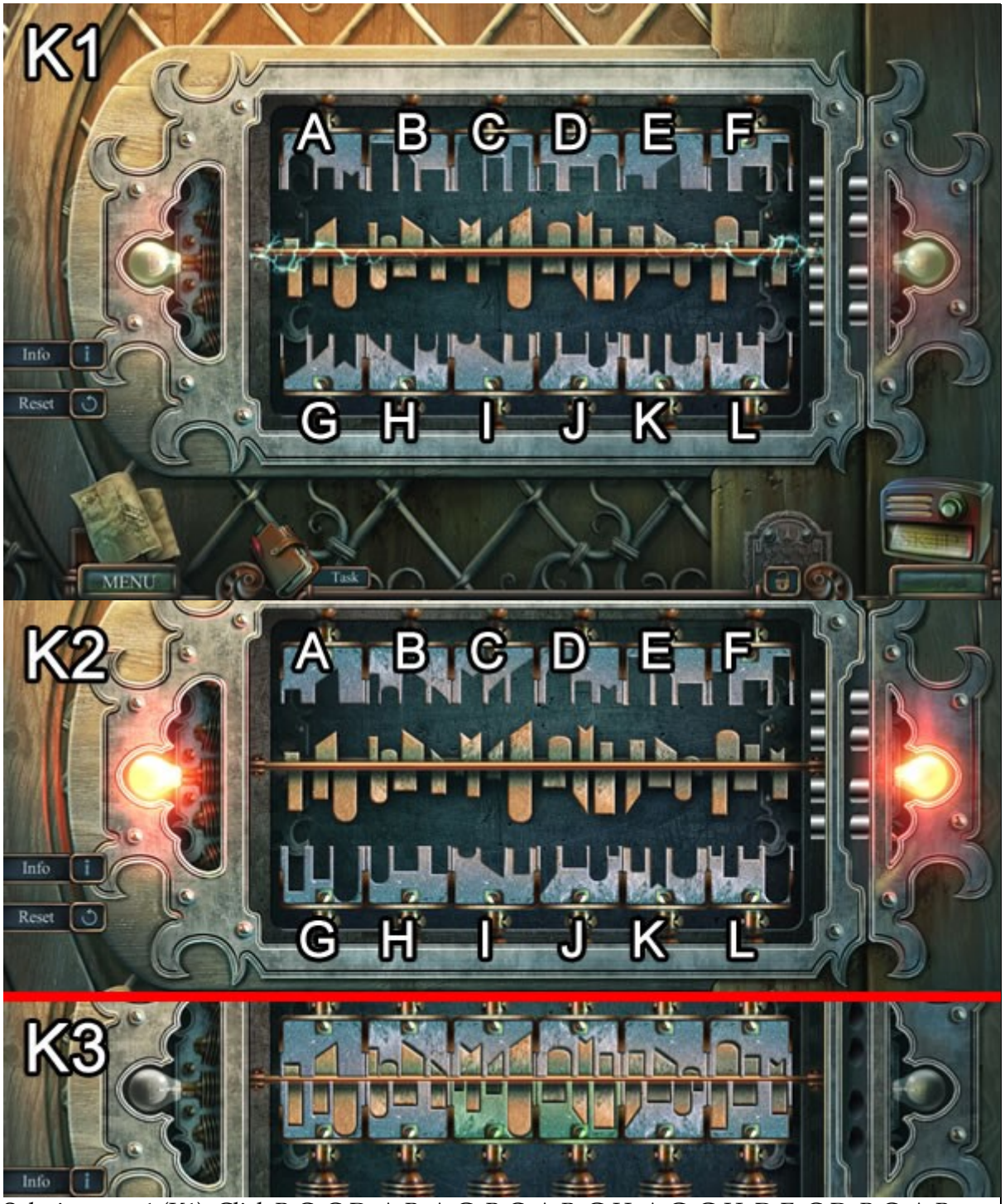
APPENDIX D: THE GROUP A GAME GUIDE

1. Click on the door. Click 3 times to clean the **debris**; Take the NOTE. — خُرد و ریز
2. Click the NOTE again in your bag
3. Take the STONE and RING; read the note.
4. Click on the old **shack**. — کلبه
5. Use the STONE and click on the window 2 times; click on the **latch**; click on the door to open. — چفت در
6. Now, try to find things in the list.
7. Click on the **hook** — قلاب
8. Now, click on the **skull**; click on the key. — جمجمه
9. Find and click the **matches**; — کبریت
10. Now, click on the **burner**; Find and click the Kettle, then, click the burner again — گاز پیک نیک

11. Click on the hot water
12. Click on the things in the picture with yellow arrow (↓)

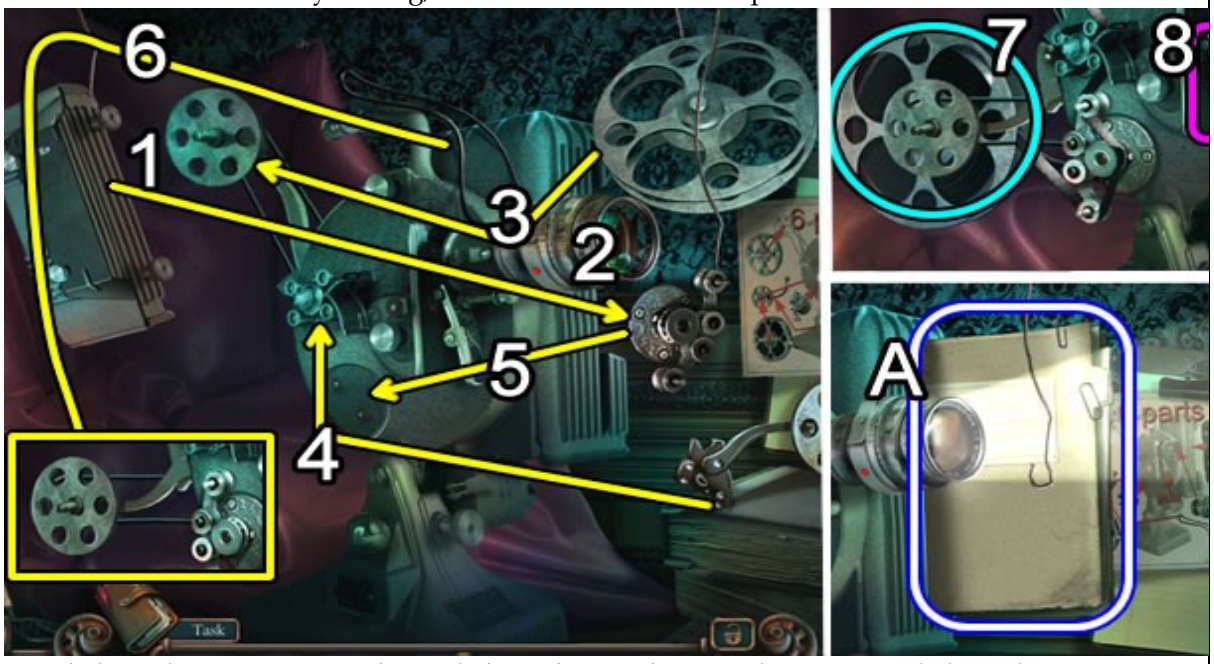


13. You get a KEY
14. Go out.
15. Click on the door, Click on the key **cover**; Click on the key in your bag and put it on the door. — سرپوش
16. The door is open, go in.
17. Click on the LADDER; click on the leaves 4 times; click on the glue
18. Go into the house. The door will close fast.
19. Click to zoom on the door. Click on the BROKEN DOOR HANDLE 1/3, 2/3
20. Click on the **carpet**. Click on the BROKEN DOOR HANDLE 3/3 — فَرش
21. Click on the GLUE in your bag; Click on the broken door handle, **fragments** stick together. — تیکه خُرده
22. Take the DOOR HANDLE
23. Click on the fixed DOOR HANDLE in your bag, click on the door hole.
24. Click on the it. There is a puzzle



Solution part 1 (K1). Click B-C, C-D, A-B, A-G, B-C, A-B, G-H, A-G, G-H, D-E, C-D, B-C, A-B.
Solution part 2 (K2). Swap K-L, J-K, I-J, J-K, F-L, E-F, K-L, F-L, D-E, E-F, K-L, J-K
Solved (K3)
25. Go inside
26. Click on the JAMES'S HAT
27. Click on JAMES'S HAT in your bag again. click the ribbon; read the note; click on the JAMES'S HAT
28. Click on the SHOVEL
29. Click on the plank on the fire, Click on the stairs to go up.
30. Click on the HAMMER, on the right wall, 3 times.

31. Go back down.
32. Click the HAMMER in your bag, click on the **portrait** — قاب عکس
33. Behind that, Click also on the PLANK 1/3 and ISAAC'S REEL
34. Go outside to the yard
35. Click on the HAMMER in your bag; click on the red **closet**, Click on the nails — گنجه
36. Click on that; it opens; Click on the sack inside
37. Click on the SHOVEL in your bag; Click on the sack to get a SCOOP OF SAND
38. Go out of the yard from the main door.
39. Click on the HAMMER in your bag, click on the door **sign**; Click on the nails and plank — علامت
40. Click on the main door, click on the house's door, you go into the house
41. Click on the SCOOP OF SAND in your bag, Click on the **embers**; Click on the PLANK 3/3 — زغال داغ
42. Click on the LADDER in your bag
43.Click on the ladder; then, Click on the 3 planks; Click on the 6 nails in your bag, Click on ladder
44. Click on the HAMMER in your bag; then Click on the LADDER 3 times.
45. Click on the French door on the left
46. Click on the LADDER in your bag; Click on the **fence-barriers** — حصار فلزی
47. Click on the 2 small doors above the French door; Click on the DIARY and the LENS
48. Click on the **drape**; Click on the projector; Look at the picture and click and drag the parts. — پرده توری
49. Click on the LENS in your bag; Click on number 2 in the picture.



50. Click on the REEL in your bag; Click on the number 8 in the picture. Click on the DOSSIER.
51. Click on DOSSIER in your bag; Click on the CODE
52. Go back to the yard
53. Click on the DIARY in your bag.
54. Upper button is C and lower button is D

55. Solution. Click C-Dx4-Cx2-Dx3-C-D-C-D-Cx2-Dx3-Cx2-Dx2
56. Read the diary; Click on the MEDALLION
57. Click on the MEDALLION and CODE in your bag. Click on the box in the yard. Code is: 1-8-9-5
58. Click on the POUCH
59. Click on the POUCH in your bag. Click on the **magnifier** and tag;
60. Now, you have a **CHAIN**
61. Go inside the house, first floor, in front of the French door.
62. Click on the CHAIN in your bag. Click on the **gears** above the French door.
63. Click on the stairs to go to the second floor, the dining room
64. Click on the pocket on the skeleton; Click on the note
65. Click on the SMALL KEY
66. Click on the SMALL KEY in your bag; Click on the **drawer**; Click on the RUST-AWAY and RUBBER GLOVE
67. Go back to the yard

ذره بین
زنجیر
دنده ها
کشو

APPENDIX E: THE GROUP B GAME GUIDE

1. Click on the door. Click 3 times to clean the **debris**; Take the NOTE.

کلبه
خُرد و ریز
چفت در

2. Click the NOTE again in your bag
3. Take the STONE and RING; read the note.

4. Click on the old **shack**.

کلبه
خُرد و ریز
چفت در

5. Use the STONE and click on the window 2 times; click on the **latch**; click on the door to open.

قلاب
خُرد و ریز

چفت در

6. Now, try to find things in the list.

7. Click on the **hook**

جمجمه
قلاب
کبریت

8. Now, click on the **skull**; click on the key.

جمجمه
قلاب
کبریت

9. Find and click the **matches**;

جمجمه
گاز پیک نیک
کبریت

10. Now, click on the **burner**; Find and click the Kettle, then, click the burner again

فَرش
کلبه
گاز پیک نیک

11. Click on the hot water
12. Click on the things in the picture with yellow arrow (↓)

13. You get a KEY
14. Go out.

15. Click on the door, Click on the key **cover**; Click on the key in your bag and put it on the door.

16. The door is open, go in.
17. Click on the LADDER; click on the leaves 4 times; click on the glue
18. Go into the house. The door will close fast.
19. Click to zoom on the door. Click on the BROKEN DOOR HANDLE 1/3, 2/3

20. Click on the **carpet**. Click on the BROKEN DOOR HANDLE 3/3

21. Click on the GLUE in your bag; Click on the broken door handle, **fragments** stick together.

22. Take the DOOR HANDLE
23. Click on the fixed DOOR HANDLE in your bag, click on the door hole.
24. Click on it. There is a puzzle

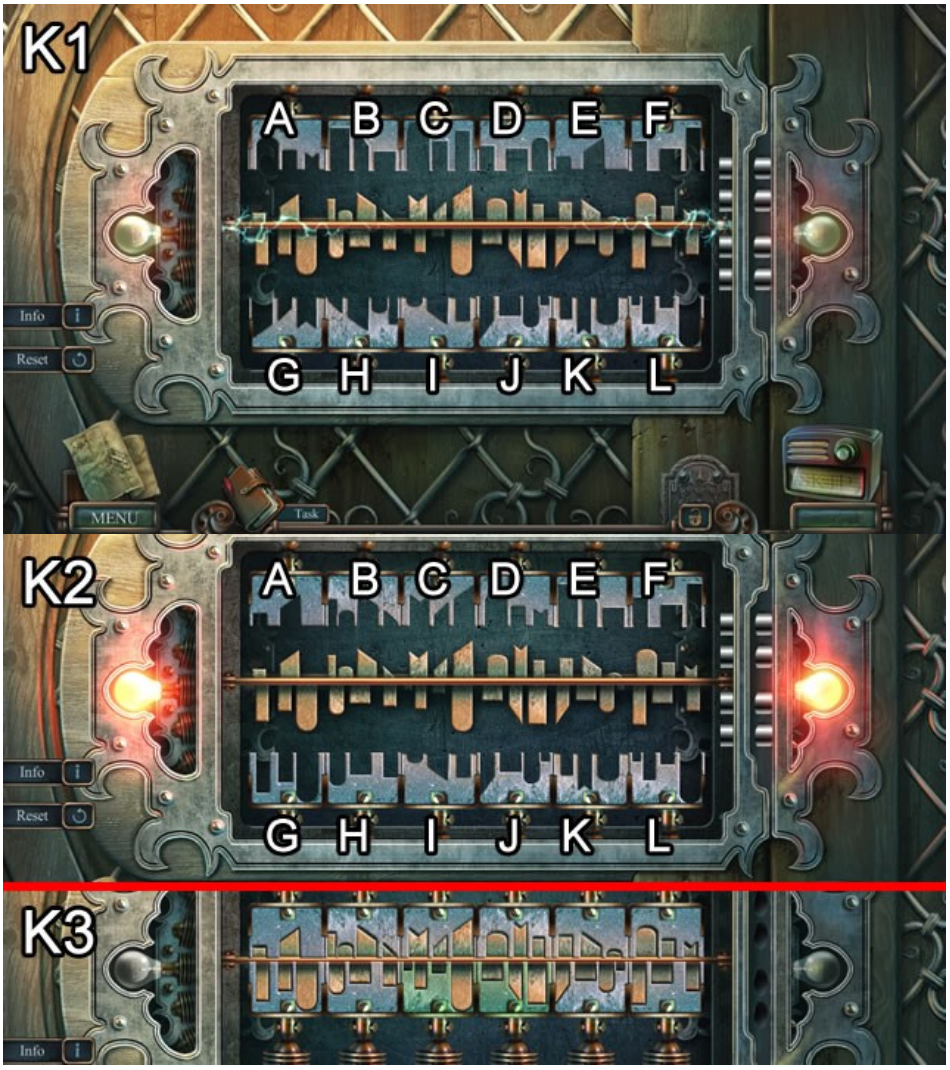
فَرش
تیکه خُرده

سرپوش

فَرش
تیکه خُرده
سرپوش

گنجه
تیکه خُرده

گاز پیک نیک



Solution part 1 (K1). Swap B-C, C-D, A-B, A-G, B-C, A-B, G-H, A-G, G-H, D-E, C-D, B-C, A-B.
Solution part 2 (K2). Swap K-L, J-K, I-J, J-K, F-L, E-F, K-L, F-L, D-E, E-F, K-L, J-K
Solved (K3)
25. Go inside
26. Click on the JAMES'S HAT
27. Click on JAMES'S HAT in your bag again. click the ribbon; read the note; click on the JAMES'S HAT
28. Click on the SHOVEL
29. Click on the plank on the fire, Click on the stairs to go up.
30. Click on the HAMMER, on the right wall, 3 times.
31. Go back down.

قاب
عکس
علامت
گنجه

32. Click the HAMMER in your bag, click on the **portrait**

33. Behind that, Click also on the PLANK 1/3 and ISAAC'S REEL
34. Go outside to the yard

قاب
عکس

35. Click on the HAMMER in your bag; click on the red **closet**, Click on the nails

علامت
گنجه

36. Click on that; it opens; Click on the sack inside
37. Click on the SHOVEL in your bag; Click on the sack to get a SCOOP OF SAND
38. Go out of the yard from the main door.

39. Click on the HAMMER in your bag, click on the door **sign**; Click on the nails and plank

زغال داغ
حصار
فلزی
علامت

40. Click on the main door, click on the house's door, you go into the house

41. Click on the SCOOP OF SAND in your bag, Click on the **embers**; Click on the PLANK 3/3

زغال داغ
حصار
فلزی
قاب
عکس

42. Click on the LADDER in your bag
43.Click on the ladder; then, Click on the 3 planks; Click on the 6 nails in your bag, Click on ladder
44. Click on the HAMMER in your bag; then Click on the LADDER 3 times.
45. Click on the French door on the left

46. Click on the LADDER in your bag; Click on the **fence-barriers**

زغال داغ
پرده توری
حصار
فلزی

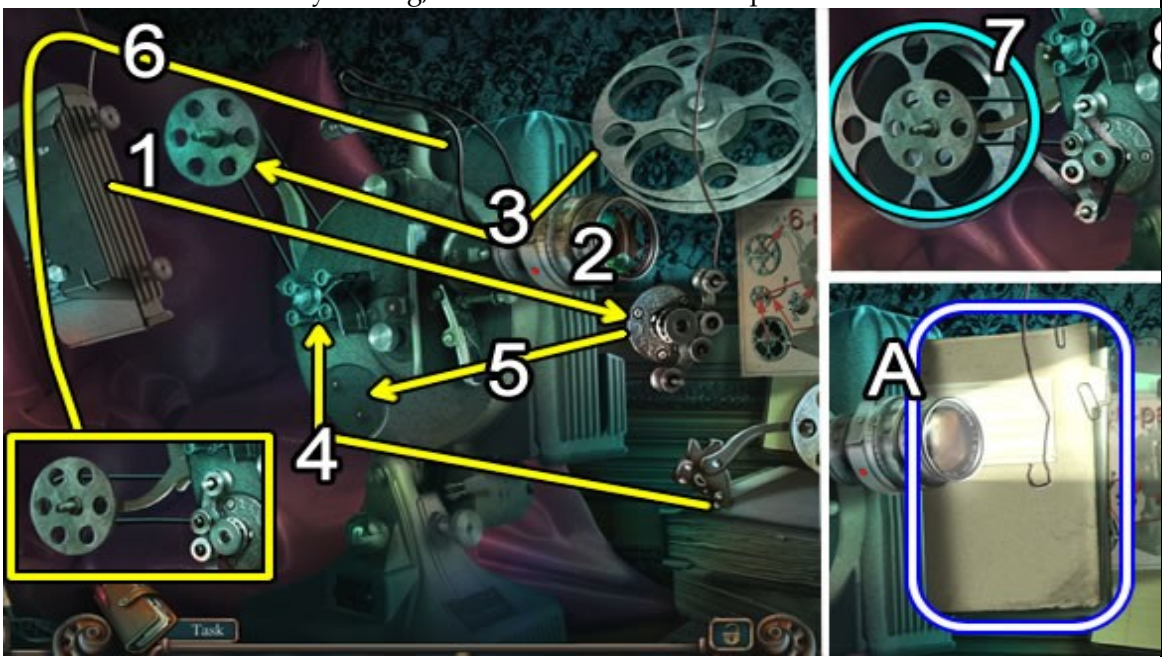
47. Click on the 2 small doors above the French door; Click on the DIARY and the LENS

48. Click on the **drape**; Click on the projector; Look at the picture and click and drag the parts.

ذره بین
پرده توری

سرپوش

49. Click on the LENS in your bag; Click on number 2 in the picture.



50. Click on the REEL in your bag; Click on the number 8 in the picture. Click on the DOSSIER.
51. Click on DOSSIER in your bag; Click on the CODE
52. Go back to the yard
53. Click on the DIARY in your bag.
54. Upper button is C and lower button is D

55. Solution. Click C-Dx4-Cx2-Dx3-C-D-C-D-Cx2-Dx3-Cx2-Dx2
56. Read the diary; Click on the MEDALLION
57. Click on the MEDALLION and CODE in your bag. Click on the box in the yard. Code is: 1-8-9-5
58. Click on the POUCH

59. Click on the POUCH in your bag. Click on the **magnifier** and tag;

ذره بین
پرده توری
زنجیر

60. Now, you have a **chain**

دنده ها
کشو
زنجیر

61. Go inside the house, first floor, in front of the French door.

62. Click on the CHAIN in your bag. Click on the **gears** above the French door.

دنده ها
کشو
زنجیر

63. Click on the stairs to go to the second floor, the dining room
64. Click on the pocket on the skeleton; Click on the note
65. Click on the SMALL KEY

66. Click on the SMALL KEY in your bag; Click on the **drawer**; Click on the RUST-AWAY and RUBBER GLOVE

دنده ها
کشو

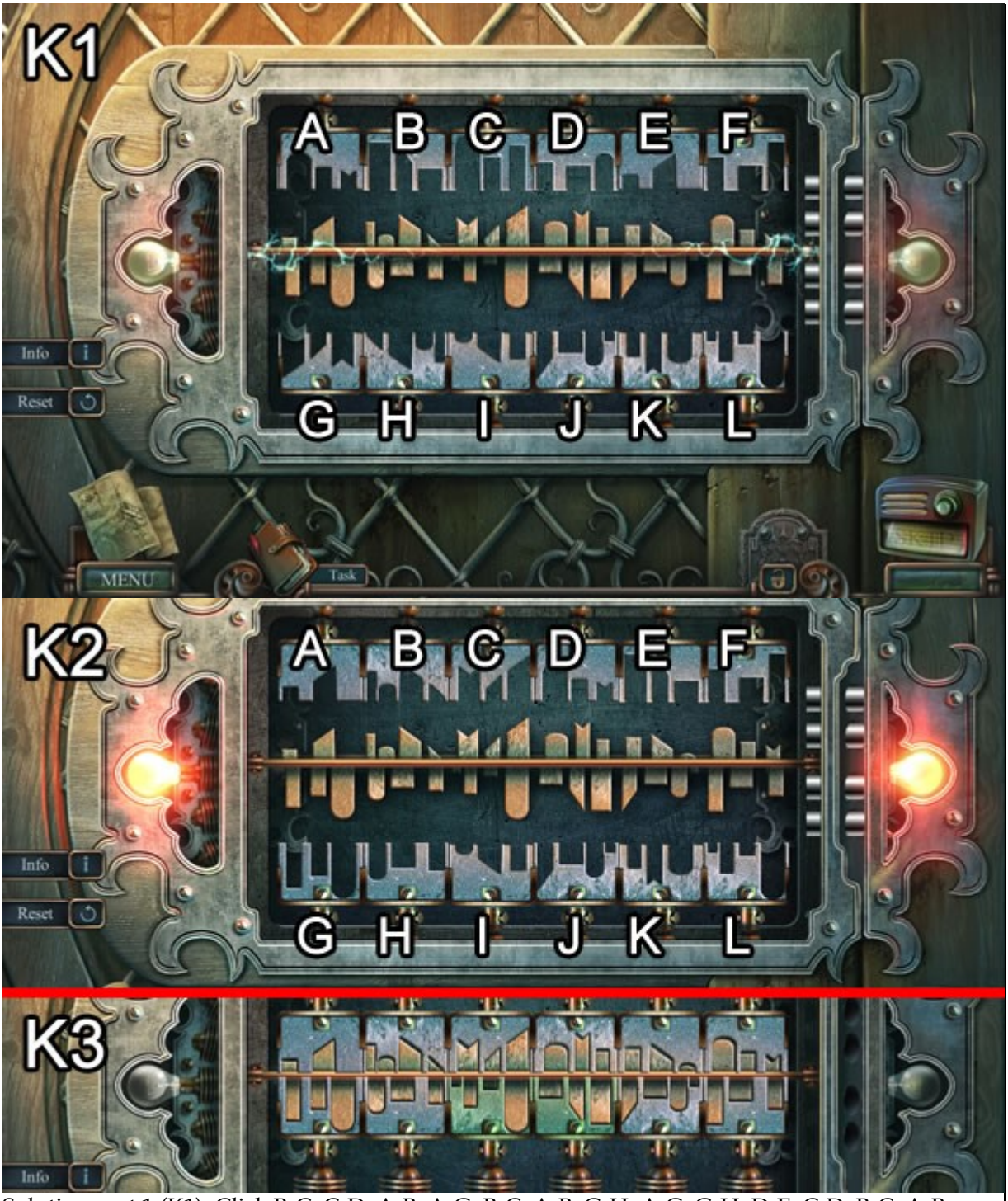
ذره بین

67. Go back to the yard

APPENDIX F: THE GROUP C GAME GUIDE

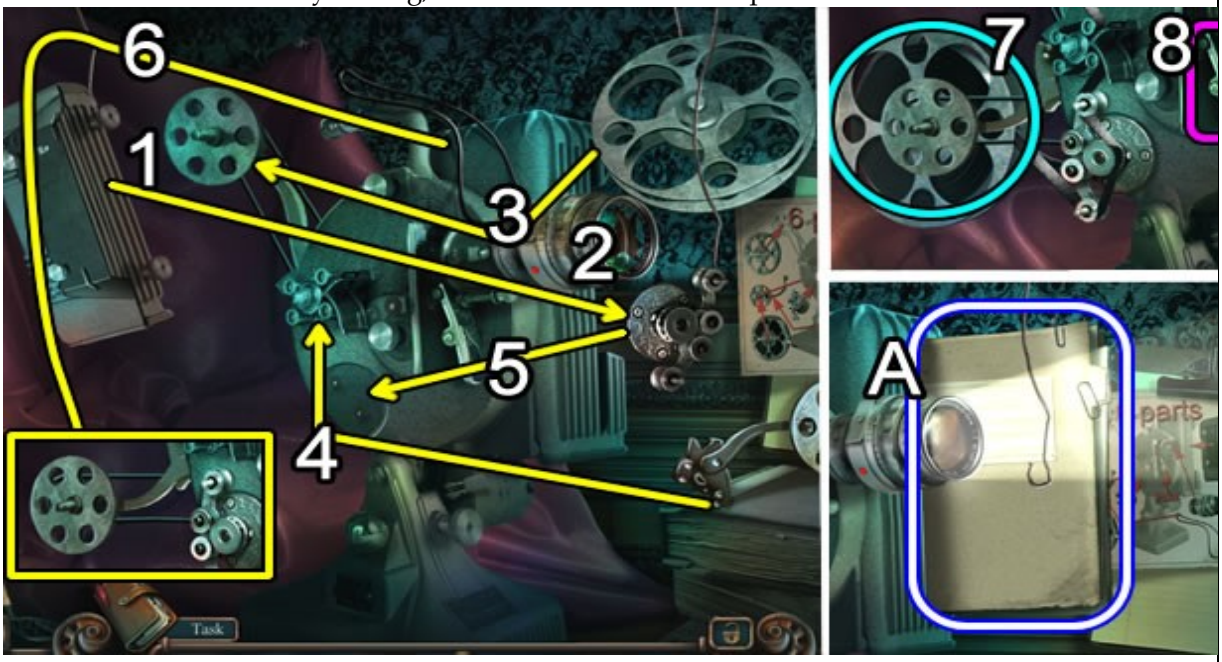
1. Click on the door. Click 3 times to clean the **debris**; Take the NOTE.
2. Click the NOTE again in your bag
3. Take the STONE and RING; read the note.
4. Click on the old **shack**.
5. Use the STONE and click on the window 2 times; click on the **latch**; click on the door to open.
6. Now, try to find things in the list.
7. Click on the **hook**
8. Now, click on the **skull**; click on the key.
9. Find and click the **matches**;
10. Now, click on the **burner**; Find and click the Kettle, then, click the burner again
11. Click on the hot water
12. Click on the things in the picture with yellow arrow (↓)

13. You get a KEY
14. Go out.
15. Click on the door, Click on the key **cover**; Click on the key in your bag and put it on the door.
16. The door is open, go in.
17. Click on the LADDER; click on the leaves 4 times; click on the glue
18. Go into the house. The door will close fast.
19. Click to zoom on the door. Click on the BROKEN DOOR HANDLE 1/3, 2/3
20. Click on the **carpet**. Click on the BROKEN DOOR HANDLE 3/3
21. Click on the GLUE in your bag; Click on the broken door handle, **fragments** stick together.
22. Take the DOOR HANDLE
23. Click on the fixed DOOR HANDLE in your bag, click on the door hole.
24. Click on it. There is a puzzle



Solution part 1 (K1). Click B-C, C-D, A-B, A-G, B-C, A-B, G-H, A-G, G-H, D-E, C-D, B-C, A-B.
Solution part 2 (K2). Swap K-L, J-K, I-J, J-K, F-L, E-F, K-L, F-L, D-E, E-F, K-L, J-K
Solved (K3)
25. Go inside
26. Click on the JAMES'S HAT
27. Click on JAMES'S HAT in your bag again. click the ribbon; read the note; click on the JAMES'S HAT
28. Click on the SHOVEL
29. Click on the plank on the fire, Click on the stairs to go up.
30. Click on the HAMMER, on the right wall, 3 times.

31. Go back down.
32. Click the HAMMER in your bag, click on the **portrait**
33. Behind that, Click also on the PLANK 1/3 and ISAAC'S REEL
34. Go outside to the yard
35. Click on the HAMMER in your bag; click on the red **closet**, Click on the nails
36. Click on that; it opens; Click on the sack inside
37. Click on the SHOVEL in your bag; Click on the sack to get a SCOOP OF SAND
38. Go out of the yard from the main door.
39. Click on the HAMMER in your bag, click on the door **sign**; Click on the nails and plank
40. Click on the main door, click on the house's door, you go into the house
41. Click on the SCOOP OF SAND in your bag, Click on the **embers**; Click on the PLANK 3/3
42. Click on the LADDER in your bag
43.Click on the ladder; then, Click on the 3 planks; Click on the 6 nails in your bag, Click on ladder
44. Click on the HAMMER in your bag; then Click on the LADDER 3 times.
45. Click on the French door on the left
46. Click on the LADDER in your bag; Click on the **fence-barriers**
47. Click on the 2 small doors above the French door; Click on the DIARY and the LENS
48. Click on the **drape**; Click on the projector; Look at the picture and click and drag the parts.
49. Click on the LENS in your bag; Click on number 2 in the picture.



50. Click on the REEL in your bag; Click on the number 8 in the picture. Click on the DOSSIER.
51. Click on DOSSIER in your bag; Click on the CODE
52. Go back to the yard
53. Click on the DIARY in your bag.
54. Upper button is C and lower button is D

55. Solution. Click C-Dx4-Cx2-Dx3-C-D-C-D-Cx2-Dx3-Cx2-Dx2
56. Read the diary; Click on the MEDALLION
57. Click on the MEDALLION and CODE in your bag. Click on the box in the yard. Code is: 1-8-9-5
58. Click on the POUCH
59. Click on the POUCH in your bag. Click on the **magnifier** and tag;
60. Now, you have a **chain**
61. Go inside the house, first floor, in front of the French door.
62. Click on the CHAIN in your bag. Click on the **gears** above the French door.
63. Click on the stairs to go to the second floor, the dining room
64. Click on the pocket on the skeleton; Click on the note
65. Click on the SMALL KEY
66. Click on the SMALL KEY in your bag; Click on the **drawer**; Click on the RUST-AWAY and RUBBER GLOVE
67. Go back to the yard

APPENDIX G: THE GROUP C WORD LIST

Debris: Example: workmen were clearing the roads of the leaves and garden debris.	خُرد و ریز
Shack: *Example: In a shack, friends gathered together for a night camp outside.*	کلبه
Latch: Example: Push the latch to open the door.	چفت در
Hook: *Example: Tom used the hook to take a box out of water.*	قلاب
Skull: *Example: He had a bad accident. He damaged his skull badly.*	جمجمه
Matches: *Example: He started a big fire with matches.*	کبریت
Burner: *Example: He used a burner for cooking eggs.*	گاز پیک نیک
Carpet: *Example: I always hide my keys under the door carpet.*	فَرش
Fragments: *Example: John fixed the glass by attaching the broken fragments together.*	تیکه خُرده
Portrait: *Example: Grandfather's portrait hangs on the wall in the living room.*	قاب عکس
Cover *Example: When it rains, I put a cover on my car*	سرپوش
Closet: *Example: My personal things are in my commode.*	گنجه
Sign: *Example: we cannot have dogs with us because "No Dog" is on the door sign.*	علامت
Ember: *Example: In the house fire, Eddy burned his hand by a falling ember*	زغال داغ
Fence-barriers: *Example: Soldiers couldn't move ahead because of poisonous fence-barriers.*	حصار فلزی
Drape: *Example: The poor man's door was covered with a drape*	پرده توری
Magnifier: *Example: Sherlock Holmes looked into his magnifier for analyzing the blood.*	ذره بین
Chain: *Example: She had a gold chain around her neck.*	زنجیر
Gears: *Example: There are a lot of gears in robots and clocks.*	دنده ها
Drawer: *Example: Maggie pulled her desk drawer out to take her pen.*	کشو

APPENDIX H: THE INTERVIEW QUESTIONS

Personal Experience & Background

- How often do you play video games? On what platform (Mobiles, Tablets, PCs, or Consoles)?
- What do you focus on when you play video game? What about this game? Why?
- Have you ever tried to learn new words from video games? Describe it.
- Are there words that you have learned from video games? Like what?
- Is there anything, in the video games, to help you understand unknown words? What about this game? Describe it.
- Do you use any tricks, when playing a video game, to help you with unknown words? What about this game?

Digging Out the Processes & Strategies

- Do you think you have learned any new words from playing this video game? For example, what?
- What do you think you have learned about the words in this game? Meanings? Pictures? Spelling? Why?
- How do you think you have learned them?
- Can you remember them easily? For example, What? Why?
- Has this video game helped you to remember / learn them easily? How?
- Did the game guide help you to learn the selected words? Why? If yes, How?
- Did it help you to remember or learn the new words or did the video game? Why?
- What did you do when you met an unknown word/selected words in the game guide of this video game?
- What did you do if you couldn't remember/recognize the meaning for the unknown words in this game guide?

Monitoring the Cognitive Load

- What was difficult about playing this video game? What bothered you the most?

How did you overcome it?

What was easy about playing this video game? What helped you the most? How?

- How is it easier to play this video game and learn new words?
- Was there anything else like the interface, mouse, keyboard, font size, font style, etc. that made it hard for you to play the game or stopped you from playing the video game easily?

Other

- Is there anything else you want to add about the game, word learning, and remembering them?

SUMMARY

Vocabulary acquisition and teaching are among the major activities in every second or foreign language classroom. However, vocabulary acquisition and teaching are challenging because acquiring vocabulary demands the acquisition of many bits of information. For instance, Ringbom (1987) sees vocabulary acquisition as a process of developing a complex interconnected matrix of knowledge systems that are accessed for both comprehension and production. He posits that when an L2 language learner wants to acquire a word, he faces different linguistic tasks such as learning the internal form (morphology), the meaning (semantics), the use of the word in a sentence (syntax), the words that it can be combined with (collocation), the words that are related to it (association), and finally, the extent to which the word can be accessed (accessibility). Thus, vocabulary acquisition is a multidimensional process. To reduce the demands and challenges that an L2 learner may face during his attempts at vocabulary acquisition, many teaching and acquisition techniques have been proposed, studied and applied. Among these, methods of incidental vocabulary acquisition have been encouraged over other methods as incidental vocabulary acquisition is considered an influential way of acquiring vocabulary from context, a motivator for extensive reading, and a promoter of deeper mental processing (Ahmad, 2001). Schmidt (1994, as cited in Laufer & Hulstijn, 2001) refers to incidental vocabulary acquisition as the unintentional acquisition of vocabulary items when learning, for instance, how to communicate.

In addition to incidental vocabulary acquisition, it had been proposed that the structure of vocabulary acquisition tasks may influence the rate and the quality of vocabulary acquisition. Hulstijn and Laufer (2001) found that vocabulary acquisition tasks that provide the opportunity for more deeply processing vocabulary items qualitatively and quantitatively assist incidental as well as intentional vocabulary acquisition. Their idea derives from the memory studies by Craik and Lockhart (1972) and Craik and Tulving (1975). In a seminal paper, Craik and Lockhart (1972) introduced the concept 'depth of processing'. They argued that "memory trace is a by-product of perceptual analysis. The persistence of that trace depends on how deep the stimulus has been analyzed" (Craik & Lockhart, 1972, p. 671). They continue that more elaboration is associated with longer persistence of the trace. In other words, "rich (qualitative) and numerous (quantitative) associations with existing knowledge […] increases the chance that the new information will be retained" (Hulstijn & Laufer, 2001, p. 541). Hulstijn and Laufer (2001) termed this notion the levels of involvement load hypothesis (ILH) The ILH offers a method for indexing the effectiveness of a vocabulary acquisition task. In this method, proposed by Hulstijn and Laufer (2001), an index of the involvement load of a vocabulary acquisition task is constructed based on three components, namely need, search, and evaluation, in which need is a motivational and search and evaluation are cognitive components. The higher the index for each component, the more effective the

vocabulary acquisition task. Although the original indexing method helped researchers in designing, knowing and predicting different vocabulary acquisition tasks, it was criticized for lack of precision (Folse, 2006; Keating, 2008; ; Martínez-Fernández, 2008; Kim, 2010; Jahangiri & Alipour, 2014; Zou, 2017). Nation and Webb (2010) later expanded the concept and introduced a new indexing checklist, based on their term, technique feature analysis, which was found to be more precise in indexing vocabulary acquisition task-induced levels of involvement load (Hu & Nassaji, 2016; Chaharlang & Farvardin, 2018; Gohar, Rahmanian, & Soleimani, 2018; Zou & Xie, 2018).

Although many vocabulary acquisition techniques have been proposed and studied, no optimal method for efficient vocabulary acquisition has yet been found. Recently, however, researchers have recognized the educational potential of digital games. In language learning, the emergence of new terms such as Digital Game-Based Language Learning (DGBLL) indicates the importance of this phenomenon for language teachers and researchers. However, the integration of digital games in language learning classrooms and tasks, especially vocabulary acquisition tasks, remains uncharted territory. Much of the existing literature on digital game and language learning is devoted to experimental studies assessing the effect of digital games on vocabulary acquisition. Meta-analyses (Chiu, Kao, & Reynolds, 2012; Chian-Wen, 2014) have substantiated claims that digital games facilitate vocabulary acquisition. A by-product of these studies has been the emergence of discussion on precisely how digital games enhance vocabulary acquisition. It has been proposed that the internal elements digital games have a cognitive-motivational effect on the vocabulary acquisition process. Digital games commonly provide rich visuals and audio that support factors known to be effective in vocabulary acquisition, such as high frequency of occurrence, variation in mode of presentation and authentic contexts (Sundqvist & Sylvén, 2012; Hwang & Wang, 2016; Zhonggen, 2018; Janebi Enayat & Haghighatpasand, 2019). Although the motivational aspect of digital games has supported vocabulary acquisition (Ebrahimzadeh & Alavi, 2016), the cognitive aspect been found to be conditional on other factors and even, sometimes, a hindrance. For example, the effect of interactivity, one of the internal elements of digital games that has cognitive effects on vocabulary recall, was found to be conditional (deHaan, Reed, & Kuwada, 2010), although, Zhonggen (2018) found that interactivity-rich games facilitated vocabulary acquisition better than less interactivity-rich digital games. deHaan, Reed, and Kuwada (2010 investigated the effect of interactivity in a musical digital game on second language vocabulary recall. They asked one group to play a musical rhythm game and another group to watch the gameplay. A t-test comparison of the results of two-week delayed vocabulary recall post-tests revealed that the watchers, who had a low level of interactivity, recalled more vocabulary items than the players. They concluded that the high level of interactivity common in digital games leads to cognitive overload and, consequently, lower vocabulary item retention and recall. This finding is important as vocabulary acquisition is a complex and multidimensional process that can be affected by many, especially cognitive,

factors. Reynolds (2017), in researching the suitability of mobile digital games for incidental vocabulary acquisition, studied internal elements in DGBL in order to find out what factors contribute to cognitive-motivation in digital game-based vocabulary learning tasks. He found that when digital games were used in vocabulary acquisition tasks, gamers adopted such components of involvement load as search, need and evaluation. Accordingly, he concluded that involvement load was an important cognitive factor in selecting a digital game for vocabulary acquisition, stating that "how useful one particular game can be for inducing vocabulary acquisition depends on the amount of task-induced involvement" (Reynolds, 2017, p. 482). Finally, he recommended the involvement load hypothesis as a precise touchstone for analyzing digital games and their suitability for DGBLL purposes.

Although Reynolds (2017) emphasized the importance of digital game-based task-induced involvement load, he did not evaluate its effect on vocabulary acquisition, leaving a gap in the digital game-based vocabulary acquisition literature. This study contributes to filling that gap by investigating the effect of different levels of involvement load (Hulstijn & Laufer, 2001) on the success and quality of vocabulary acquisition tasks performed using a digital game. The research questions were as follows:

1. What is the effect of the digital game, in different levels of involvement load, on the acquisition of target vocabulary items?
2. Which dimension and scope of word knowledge, either receptive (recall/recognition) or productive (recall/recognition), are acquired significantly better after completing digital game tasks in different levels of involvement load?
3. Does interacting with the digital game tasks, in different levels of involvement load, make significant differences in vocabulary acquisition?

To achieve the aim of this study, 30 Persian speakers (14 males, and 16 females, 13 – 15 years old) were randomly recruited. Their homogeneity was evaluated by the *Vocabulary Size Test.* It was found that selected participants had a knowledge range of 2700 to 3000 word families. This study followed a semi-experimental design. In other words, there was a pre-test, treatment, and post-test. After sitting for both the pre-test and *Vocabulary Size Test,* participants were randomly assigned to three groups, tagged, A, B, and C. Each group comprised 10 participants who were asked to play a digital game in pairs. Two pairs were randomly selected from each group for the purpose of collecting concurrent think-aloud data. The digital game tasks for each group were designed so that group A's task induced the lowest (index 7), group B's task a moderate (index 9) and group C's task the highest level of involvement load (index 12). The task-induced levels of involvement load were measured using technique feature analysis (Nation & Webb, 2010). After the pre-test, the concurrent-think-aloud participants attended a warm-up session to learn how to think aloud. Next, all participants performed their main task, playing the adventure commercial digital game *Haunted Hotel: Death Sentence* in pairs. They were instructed to read a game guide tailored to their group task and solve the murder mystery in the game. The

game guide foregrounded 20 target words. The target words were the names of inanimate objects, lexically nouns, knowledge of which was necessary for completing the game. The target words were not repeated more than twice in the game guide. Three weeks after they had finished played the digital game, they sat a delayed post-test. The concurrent-think-aloud participants were interviewed immediately after completing their task.

Collecting both qualitative and quantitative data provided an opportunity to investigate more precisely the effect of digital game-based vocabulary tasks on the acquisition of the target words. The quantitative data analysis showed that although the digital game-based vocabulary acquisition tasks were effective in the acquisition of different dimensions and scopes of a word, i.e. productive and receptive recall and recognition, the order of learning was as reported in previous studies. In this study, participants scored higher on all scopes of productive knowledge (mean rank = 17.33) than receptive knowledge (mean rank = 17.03) of the target words. Moreover, although group C, with the highest level of involvement load, outperformed other two groups in the post-tests, group B, with the moderate involvement load, showed the poorest performance, thus contradicting both the ILH and previous findings. The qualitative data analysis showed that, in multimedia contexts like digital game-based tasks, factors such as task structure, context, and strategy selection can play a crucial role in vocabulary acquisition. Accordingly, each group followed a distinct learning approach that included universal moves and exclusive strategies in performing their digital game tasks. For example, universal moves, i.e. moves common to all three groups, included *information search, negotiation, turn taking,* and *trial-and-error*. Exclusive strategies applied by group A included *input enhancement strategies,* exclusive strategies applied by group B *inferencing* and *hypothesis testing strategies,* and exclusive strategies applied by group C *memory search, feedback request* and *word association strategies*, and a group-specific move termed *planning*.

Based on the differences in learning approaches, it was speculated that these factors might have caused the poor performance of group B in this study. Group B's task encouraged the participants to employ inferencing techniques. Inferring from context invokes distinctive processes such as decision-making generators, or processes for deciding which definition to use, and evaluators, or, hypothesis testing, which test the selected definition (Nassaji, 2003). Although these processes have been found to support vocabulary acquisition (Hulstijn, Hollander, & Greidanus, 1996), they do not explicitly initiate form-meaning links. Furthermore, due to multiple alternatives in the task of inferring meaning from glosses, these processes might have been counter-productive and led to confusion (Martínez-Fernández, 2008; Bao, 2015). Thus, as the think-aloud data analysis seems to suggest, participants may have been puzzled by the many alternatives, causing them to feel lost in their decision-making and hypothesis-testing. As a result, they were unable to invoke the form-meaning link.

These findings indicate that prospective language teachers must be aware that integrating digital games in language-learning courses, while possible, must be done with care. They must know that the pre-teaching of target words,

provision of a game guide, possibility of peer interaction, and tasks that induce a high level of involvement load can enhance both quality and quantity in target word acquisition and retention. Furthermore, they must be aware that the genre of the digital game to be used is also important. The adventure genre is highly recommended over other genres for boosting vocabulary acquisition due to its functional mixture of gameplay and story. Thus, teachers can ensure that both the motivational and mental requirements for high quality task performance and efficient vocabulary acquisition are met. On the other hand, teachers should perhaps avoid inferencing techniques in digital game tasks, as these strategies not only do not assist vocabulary acquisition but also confuse learners.