

Utrecht University

Department of Information and Computing Sciences
Game and Media Technology - Master Thesis

Assessment of the effects of challenge in the semantic decoding ability in a Videogame

Author

Pedro Fonseca Folhadela Lopes de Moura

Supervisor

Dr. Remco Veltkamp

Second Examiner

Dr. Christof van Nimwegen

Daily Supervisor

Dr. Herre van Oostendorp

ICA - 5795176

March 14, 2019



Universiteit Utrecht

Abstract

This study was performed in a Portuguese school, for students of age eleven, to research the effects of competition in the semantic decoding ability of the students. More specifically, using a video-game as a tool of motivation, BattleQuiz (a quiz style game, from Liftov). Comparing the results from two groups of students, that played the game with and without the influence of competition, with a third control group throughout two phases of different questionnaire games. No significant relationship was found between semantic Decoding and Competition, therefore it would be necessary to perform further research in order to take deeper conclusions on this matter. Due to this fact, the study became a pilot study and the need for alterations for a future study was registered in this paper together with the development of a new game app that will serve the next study.

Acknowledgements

I would like to thank my supervisors dr. Remco Veltkamp and dr. Herre van Oostendorp for their patience and endurance in keeping me on the right path and not giving up. I would also like to thank Jaap Gerretsen and Antoon Sturkenboom for their eagerness in helping and granting the study access to their company game. Lastly, I wish to thank my friends and family who were tireless in their support and encouragement throughout this moment of my life.

Index

Introduction	3
Context	3
Semantic Decoding	3
Memory retrieval	4
Learning	5
Elaboration	6
Competition	6
Main Research questions	7
Hypothesis	8
Method	8
Equipment	8
Game: Battlequiz Application	8
Interface: Creating and using Questions and Answers	9
School Setup	10
Material	11
Session Setup	11
Groups	11
Competition Stimuli	12
Phase Preparation	13
Tests	14
Dependent Variables	14
Variable Recording	14
Procedures	14
No Competition Group	14
Competition Group	14
Control Group	15
Results	16
Summary	17
Discussion	19
Conclusions	19
Reference	20
Tables	22
Appendix	23
Appendix 1 - Phase 1 Questions	23
Appendix 2 - Phase 2 Questions	28
Appendix 3 - Future sentence example	30
Appendix 4 - Developed Game App	31
IDE's	31
RunTime Environments	31
Code architecture files	31
Start Up	31
Browser running and testing	31

Introduction

The project aims to use a quiz game application (BattleQuiz by Liftov) to assess if the competition can enhance the learning experience of semantic decoding for children of the age when this it is mostly developed.

Even though games have been shown to have positive influence in the learning process of knowledge, it is not yet proven if they can affect the learning of more abstract concepts, such as inferring and deducing (characterized as Semantic Decoding in this thesis).

Thus, the main research question in this thesis is: Can the semantic decoding ability be influenced by playing a game, particularly focusing on the presence of competition?

This assessment is done by having two phases of testing on three groups of students of age around eleven. In the first phase, each group will be in different conditions, while answering the first set of questions. The first two groups will play game application, one of which will have the presence of competition and the other, lack thereof. The third group will be a Control Group, which will be presented to the same ideas used in the game by the others two groups. This will be done by answering the same questions in paper, having no access to the game interface or competition factors.

The second phase will be the same for the three initial groups, all of them will play the game without the presence of competition and being presented with a second set of different questions.

By recording the speed (reaction time) and correctness (score) of each participant for each of the phases we will be able to compare and possibly relate the different conditions.

Context

Since this study will be regarding reading ability, it is important to understand many of the conventions and findings regarding how learning and knowledge retrieval work, as well as how reading is learned specifically.

This study is based on theory of the following books. For the learning and general knowledge retrieval theory, the book “Human Cognition: Learning Understanding, and Remembering” written by John D. Bransford (1979) was studied [1]. In regards to the process of learning to read and what that implies for a student (of the age tested on this study), many of the conventions that will be mentioned in this thesis were taken from the book “Aprender a ler, da aprendizagem informal à aprendizagem formal”, which translates to “Learning to read, from informal learning to formal learning”, by Prof. Fernanda L. Viana and Maria M. Teixeira [2].

Semantic Decoding

For the experiments that will be performed in this study, semantic decoding is an important concept to address. Its presence in education has been studied around the world with the keywords: **Effective elaboration, semantic representations and categorization.**

Decoding is defined by translating non familiar signs into a familiar code. Semantic Decoding refers to Identifying words and translating their meaning.

Reading is composed of the Ideographic code, and the graphical-phonetic code. The Ideographic code establishes a connection between the written word and the concept (Dechant & Smith 1977) [3]. The graphical-phonetic code establishes a connection between a word and a sound (Lobrot 1980) [4]. Learning takes both codes into account, but ideographic code is the prime focus for this study, as there are no sounds in it.

In the book “Ler e ensinar a Ler”, translated to “To Read and teaching to read”, conventions on the teaching techniques for reading in Portuguese schools are addressed (by Inês Sim-Sim, Cláudia Ramos, Eveline Silva, Manuela Micaelo, Maria M. Santos and Palmira Rodrigues, 2006) [5]. It is given outstanding importance to the speed and precision of decoding, as they determine understanding (Lesgold et Al., 1985) [6]. Stating how concentration and attention are not the same concept. Findings show that, in reading, more concentration leads to less attention which in turn results in less understanding. The more the reader is concentrated in the decoding of the signs for each word, the less he is able to give attention to the meaning behind many words together. More precisely, quick readers, are able to digest more words and the meaning behind them. This is according to the reading process model proposed by LaBerge D. & Samuels (1985) [7].

Decoding speed of a word can also be respective to the category in which they are placed in terms of how common they are, also known as *cue validity* of a word. In the paper “Basic Objects in Natural Categories”, it is shown how different words are categorized in the same category word, be that in adding complexity or removing it (Eleanor Rosch, 1976) [8]. For example, a “Chair” is considered a basic object, on the other hand, “Furniture” encompasses chairs and others, so it is considered a Superordinate category, while in the contrary, “Kitchen chair” is of a Subordinate category. This taxonomy helps determine *cue validity*. Basic level words are shown to be the first known by children, and still the most used by adults. This is why they are considered to be of the highest *cue validity*. They are also defined to be the most definable by means of visual perception and by sensory motor interactions with an object (leaning back on the imagery concept for learning). The level of the word is not the same for everyone, it is affected, just like elaboration (referenced next), by the extent of knowledge of the reader.

Memory retrieval

Since this study focuses on the retention of an ability, such as the decoding ability, it is important to understand how memory is accessed. To learn information is to hold it in memory and to be able to recall it when necessary. The recall ability we possess is influenced by many variables. From the tests performed in the paper from Morris, Bransford and Franks (1977), triggers were used to stimulate memory, as to provoke recall of relevant information. It was found that, superficial triggers and triggers that stimulated semantic processes have variable results [9].

Superficial triggers are those without depth semantically, in the fillable sentence“(blank) rhymes with log.”, the answer is “dog”, but the grammatical logic is not entirely connected to the concept of a “dog”. On the other hand, triggers that go more in depth, such as the fillable sentence “(blank) has ears.” make more sense to have the answer “dog”, the concept of a dog calls to it having ears.

If the test performed is appropriate to semantic processes, it harbors better results in that sense. Since the tests performed in this study were created to appeal to the semantic process, which is the semantical decoding and the speed of the decoding, they will allow the participants to improve the recalling ability.

Another important variable is how the organization of retrieving information is done. When stimuli is done consistently and relevantly with the appropriate word cues, it shows to have a positive effect on recall [10]. Could this be an important component of the context knowledge? Different cultures will add different meaning to words, in high frequency and consistency, therefore, would this be relevant to how semantical decoding is processed? As it shall be clear in the coming chapters, context knowledge is very relevant to semantical decoding and elaboration. One could argue that, with high

frequency of learning situations, good learners develop. Good learners are shown to develop organizational structures for learning, this has suggested that the development of organizing structures can have a powerful effect on retention of knowledge [10].

Retention (or memory) is considered different from the ability recall (or retrieval), even though the latter is always used when testing memory. The retrieval-cue effectiveness of the stimuli (such as trigger words) was shown to be mostly dependent on the learning activity performed. Giving even more importance to the activity, which poses relevance to this study, which has total focus on the learning activity, method and motivation [11][12][13][14][15].

Cue recall accuracy increases with the use of congruous and applicable words to the triggering sentence [12]. It also earns positive results when there is an increase in the elaborateness (further developed in the next section) and complexity, “retention may be strongly influenced by the degree to which elaborative phrases help people use what they know to understand more precisely the significance of particular concepts;” (page 79, chapter 3, section 3 of the book [1]). Contrary to it, forgetfulness, occurs when there is lack of relevant information presented. These concepts are further expanded in the elaboration section, more specifically, the precise elaboration section of this thesis.

Learning

Learning (recall of knowledge and skills) is influenced by three major variables. Meaningfulness (semantical value), imagery and frequency of occurrence.

Meaningfulness (identifiability) is respective to the semantical value of a stimulus (such as a word that triggers an idea or concept). It was also found that more semantical value, earned faster response times in testing.

The second mentioned variable is imagery, it is the ability to create images of words and how it affects how easily they can be learned. Allan Paivio (1971) showed that images are more easily retrieved and serve a conceptual references to which response time can be attached [16]. Overall, words with high levels of imagery produced better cued recall, this was verified when pairing words of low and high levels of imagery in different combinations. Additionally, the author also proposes a dual-code theory, separating the two interconnected memory systems, verbal and visual (Paivio, 1971, 1975, 1976) [16][17][18]. A word like “alligator” is stored on both verbal and visual, making a significant enhancement of retention.

Imagery and the the dual-code theory were used in the questions and answer creation, using easily recognizable visual metaphors, for example, words such as “submarine” and a “bat”, are very imaginable, both verbally and visually. As such, these words are actually used on questions from this study.

Lastly, the variable frequency of occurrence; could it be tied to previously referred organization of retrieval? Once again being possibly related to contextual knowledge? Counter to expectations, higher frequency does not make the stimulus more recognizable. Actually, lower occurrence, makes the word more distinct (contrasting), and therefore, noticeable and recognisable.

It was also shown that, the weight of imagery and frequency of occurrence have more significant effects than meaningfulness. Which shows that this study could have difficulty in earning obvious and conclusive results, as it is heavily reliant on semantics.

Elaboration

Elaboration is defined by the ability to find, infer or create, meaning and additional non literally observed concepts.

In the paper, “Constraints on Effective Elaboration”, it is expressed that, for reliable retention of new abilities, it is important to take into consideration the knowledge of the student in the addressed domain, and the practical use of the newfound knowledge (Barry S. Stein and John D. Bransford, 1979) [19]. For example, if a student is knowledgeable of marine life, because of his particular upbringing at home, he will find advanced information on that subject to be of average difficulty to understand, but a student without such a background would find it very difficult. This competence and difficulty in a field also affects the motivation, as it will be further expanded later in this paper. Adding to this, retention is also improved when the new information shows to be useful, so redundant information can actually be in the way of the process.

In the paper of “Elaboration and Knowledge Acquisition”, it is stated how crucial the role of domain related knowledge is for knowledge retention (Barry S. Stein and John Bransford, 1984) [20]. Elaboration can be used to facilitate retention acquisition of additional expertise. In their study, it was shown that, elaboration that was relevant to the topic, had a very positive effect on knowledge retention; as mentioned in Memory Retrieval section of this thesis. More specifically, as long as the concept in the sentence was precisely justified, the relationship between ideas present in the sentence had no significant impact on retention. Could it be used in training to improve skills? Could it have negative effects, if the new information is not relevant to the topic? Hopefully this study will give some insight into this.

Precise Elaboration

As to develop a specific ability or concept, the elaboration tests used in this study are that of precise elaboration. Which is defined as being presented with situations in which the target concepts (“tall”) are relevant or non arbitrary (“reaching the top shelf”), for example: With the

base sentence “the tall man buys the crackers...”, the precise elaboration would be, “...that were on the top shelf”. Many of the experiments and tests performed to measure elaboration, end up being done with precise elaboration.

In the papers “Constraints on effective pictorial and verbal elaboration”, explanatory elaborations (precise elaboration) had significantly higher results of recall over arbitrary elaborations, both when using pictures or not as the stimulus (by Barry S. Stein, 1987) [21]. For this study, images will not be implemented as part of it, as this option was not available in the used game app (and could not be implemented). Some of the sentences in the appendix of the referred paper were used as a basis for some of the sentences, questions and answers in this study. Other sentences that were also used as a basis were from the paper “Elaboration and knowledge acquisition” [20].

Competition

Challenge is also being introduced in this study, as to improve the engagement of students, as it is a big influence in motivation, consequently, being very beneficial for the study (SCI Model, Ermi & Mäyrä, 2005) [22]. In the paper of “A Comparison of lecture-based and challenge-based learning in a workplace setting”, results showed that more interaction, due to challenge based-learning, had students perform better than lecture based learning (Timothy K. O’Mahony, Nancy J. Vye, John D. Bransford, and Elizabeth A. Sanders, 2011) [23]. More interactivity between the participants, gives place to more verbalization of concepts and their interconnections, which further develops their understanding.

In the paper of “Who but not where: The effect of social play on immersion in digital games” it is shown that there were some misconceptions on the effect of social play, which encloses competition (Paul Cairn, Anna L. Cox, Matthew Day, Hayley Martin, Thomas Perryman, 2013)

[24]. While before, social play was indicative of “fun”, which is, of itself, positive for the motivation of players, it was believed to have negative effects on immersion.

Immersion is currently the “buzz” word in game development and design, as the main motivator. It is widely studied and categorized, as mentioned before (Ermi & Mäyrä, 2005) [22]. Ermi & Mäyrä reveal that competition actually has a positive effect on immersion. Games in education have as main tool, the abuse of the inherent concept of video games, motivation. So having competition stimulating both “fun” and immersion makes it a significant factor for motivation. So it is relevant as a tool for any educational game.

In regards to early education and human development many conventions were taken from the book “Educational Psychology: A Development Approach”, by Norman A. Sprinthall, Richard C. Sprinthall [25]. Competition has been shown to be a cultural construct instead of instinctive one, but since this study and research is applied in European countries and not in the Zuni tribe, where cooperation is paramount, compared to competition, it can be taken as a given, (“Towards a Theory of Instruction”, Bruner J. S., 1966) [26].

In early development, competition does not tend to develop positive traits in children, but not all personalities have synergy with purely cooperative learning situations. The simple presence of a group stimulates feelings of anxiety and generates situations of competition (this will be present in the study as the experiments are made in group). Still, there is a place for competition in helping in low motivation experiences, as it elevates fulfillment. Adding competition and anxiety can make small and simple tasks have a greater feeling of fulfillment. This works especially in low motivation situations, such as a dealing with a subject in

which the child is less confident and because of that, less motivated (“The effects of cooperation and competition upon group process” by Deutch M., 1949) [27].

The confidence in a subject is connected to the motivation and interest of a student in said specific subject. So students that have low competence to some subjects usually lack any interest or motivation to learn them. To battle this, it is encouraged to have smaller simpler challenges of the subject, as this gives an easier sense of competence to the student, motivating him/her (“Motivation reconsidered: The concept of competence” by White R., 1959) [28].

Still, competition shows to dominate the path to success in the future (“Some descriptive aspects of two-person non-zero-sum games” by Minas J. S., Scodel A., Marlowe D. and Rawson H. , 1960) [29]. Still, there are positive and negative effects, as stated in the paper “How competition and heterogeneous collaboration interact in prevocational game-based mathematics education” (by Judith ter Vrugte, Ton de Jong, Sylke Vandercruysse, Pieter Wouters, Herre van Oostendorp, Jan Elen, 2015), competition can have positive effects on the collaboration of for above average students, but the contrary reaction from under average students [30]. Though this was only applied in the context of a mathematics game.

Main Research questions

What is the effect of the presence of competition in teaching abstract concepts such as elaboration and the decoding speed? After recording the data from the game being played, it was tested by performing a one-way ANOVA between the three groups (focusing on the differences with the Competition Group).

What is the effect of using a videogame tool for teaching skills of elaboration and decoding speed?

This was tested by performing a one-way ANOVA between the three groups (focusing on the difference with the Control Group).

Hypothesis

The hypothesis and past studies point towards having the group with competition to have the greatest influence and gain, followed by the group without competition and then the Control Group as described in this study.

A one-way ANOVA for the three groups in each phase will be performed, as to compare their means and find their significant differences. It is expected that, competition will have a positive influence in scoring and reaction time, and no competition (whilst still playing the game) will also earn better results than the Control Group (which does not play the game in the first phase).

Method

In this section we will address the method used in the study, divided in Equipment, Session setup, Tests and Procedure for said tests.

Equipment

The following section is organized in three parts: Game, the base game by Liftov, the Interface, presenting the reasoning behind the questionnaires created and the School Setup, which addresses the school and participant test logistics and influence on said test.

Game: Battlequiz Application

In this section it is described how the game (by Liftov) was suited to the experiment and how the experiment was affected by the game's limitations. The game was an application used for training employees of companies with specific knowledge. The employees play battles (quizzes) by challenging each other, and are ranked in each

arena (area of knowledge of questionnaire). The analytics are automatically collected by the application's online server in the form of averages per unique question answered.

Constraints

For each game played, seven questions are selected from a chosen set of questions according to the chosen arena, this number cannot be changed. For each of the questions, four answers exist and one will be correct, this cannot be changed. Pictures connected to each questions cannot be added or implemented.

Gameplay

Participants will play the game by either challenging the computer or another player to a battle multiple times. Playing , in each battle, through quiz type game, answering each of the seven questions as correctly and fast as possible.

Interface: Creating and using Questions and Answers

For the experiment, the Forced-Choice method of questions and answers was used. This was due to both the limitations of the Battle Quiz customizability and to the advantages inherent to the method. The Forced-choice method implies that an amount of possible answers are given with the questions, and that players have to accept one of the options, and cannot abstain [1].

Affecting Recall

Questions and sentences in this study test the elaboration speed and ability of players. But to understand the questions or concepts being given, it is important to be able to remember and recall some of the content in the sentences and answers. So, taking into account concepts that affect recall is important.

In Section 2, Chapter 4 of the referenced book of Human Cognition, the structure of sentences is proven to be important for the recall ability [1]. Sentences that are unintuitive, either linguistically

or verbally confusing, have bad structure. One example would be the sentence “That that is not that that is not is that it is”. How a concept is used in the triggering sentence also affects its structure, as it affects its potential as a retrieval cue. Adding to this, since children have a tougher time reading, it is important keep the size and complexity of sentences appropriate.

Sets of Sentences

It is emphasized how learning can be affected by not only the different components of a single sentence, but also by the relationship between different sentences (Section 3, from the book [1]). An example is how upon reading a book narrative, one grows to understand a character’s multi dimensionalities better throughout the book, than from reading a single sentence. The findings suggest that people acquire something more than the mere list of individual sentences given. Different perspectives and approaches (sentences) on the same concepts (that are non arbitrary) affect our understanding of their meaning.

These intersentential (between sentence) relationships affect greatly the ability to perform the study. On the mentioned book, it is stated how the speed of comprehension is dependent on these intersentential relationships (reaction speed)[1]. In the test described on the mentioned book, sentences of higher precision earned higher recall results than those of lower precision; the sentences created in the study of this thesis are all modeled with high precision. On higher precision sentences, participants need to recall concepts from the words in the sentences, coupling them with semantically relevant information to choose the right choice. Though, the questions created are for the general ability of semantic decoding, it is hard to really point to one specific concept.

Learning Dependencies

Studies shows that learning is heavily dependent on the learning activity itself. For this study

specifically, Integrative abstraction, which, as stated before, refers to how the acquired knowledge equates to more than the sum of its parts (elaboration). Pott’s results (1972, 1974) also show that participants verifying the truth of statement were quicker when inferring without previous knowledge than to verifying with previous knowledge [31][32]. Pointing towards how people construct integrated structures rather than simple sentence memory representations.

Though the questions created are not coherent to one concept, there are general themes in the questions regarding to the academic knowledge goals and areas of the participating student’s year. These topics also points towards another study’s findings described in the book [1]. In addition to variables like topical information, the use of relevant examples can also have a powerful effect on people’s abilities to learn from textual information.” [33].

As we have seen, learning can be eased for the general population by individual word cues, sentence structure and the semantical relationship between sentences.

Now we can focus on the relevant details for teaching children. The 3 main variables that affect relevancy (Klare, 1963) are [34]:

- Word difficulty, usually based in the frequency (Thorndike and Lorge, 1944) with which children are presented with some words, in turn leaning towards the category of Base Level words (as mentioned before) [35]. In this study, questions and answers use simple and recognizable words and have a high probability of repeating themselves.
- Sentence Length, longer sentences add to complexity. It is important to avoid conjunctions sentences but over-simplification has a negative effect. Oversimplifying makes it difficult for children to read the subtext (as they do not relate as easily to the context they are

reading). This simplification can make some related information seem only arbitrary if not further exposed.

- Sentence complexity, a moderate degree of synaptic complexity is proven to facilitate comprehension. Children prefer to read what is moderately complex and to look for an answer over low complexity sentences.

Some authors state that children have difficulty dealing with recalling the chronological order of events. But this is not totally correct, it is important to structure the story and events appropriately to the age of reader.

School Setup:

Participants

To analyse the growth of decoding ability, participants needed to be around the age when such concepts are developing, specifically around eleven years of age. So, a school was needed to provide the opportunity of finding participants. It was also decided that, to make the study more feasible, the goal number of participants would be 36 students, twelve for each of the three group. In this section, we go through the necessary logistics needed for the school's participation in the study and subsequent participants' approval.

School agreement

For the experiment, it was important to reach out to a school that had students in the target age of eleven. In this case, the school Colégio Nossa Senhora de Lourdes, replied with availability. This school provides classes from kindergarten (four months and a half) to the end of middle school (age fifteen). Still, because of legal rules, it was necessary to have the informed consent of the parents for the participation of each student (child).

Facility and devices

For the testing hours, the school provided a computer room with access to 20 computers (with tactile screens), which was used to login to the browser application and play the game with all the participants of each group at the same time.

Student's guardian legal consent

Students were reached out to participate in the testing. The Master Student responsible for the study went to each class (portuguese fifth and sixth grade, ages ten to twelve) to provide a small introduction in the relevance to the study, motivating the students to ask their parents for permission. Not revealing any specifics about the study, except for its objective (to apply games in education).

In the end there were 28 confirmed participants. Participants were arranged in the three groups, considering only gender and grade. Parents were then emailed as to be informed of the specific study schedule in which their child would be participating, according to the group they were assigned to. This division is further expanded on the **Tests** section for each group.

Dates

The testing was agreed to be performed, between the end of May and the beginning of June. As classes officially end in the 23rd of June and some students (from the fifth grade) would be having national benchmarking exams, the specific dates for the tests would still be agreed at a later date.

It was also decided that the tests would be performed out of schools hours, as this collaboration was agreed during May and April, which is very near the end of the School Year, resulting in less schedule flexibility. The testing sessions were performed at the end of the lunch break, before the start of classes, from 14:00 to 14:30.

Material

To determine and evaluate the skill, that is semantic decoding in young participants, it is necessary to take into consideration the concepts of decoding and elaboration.

Since the game (from BattleQuiz) is a trivia game, it was decided that the study design would be made with questions or sentences, and multiple answers. Due to some miscommunication, the amount of answers was different than it was earlier expected. The game was tested, having four possible answers (relative to each question), and it should have been two static answers instead, being: “important” and “not important”. Though, this was a constraint of BattleQuiz as well (always has 4 answers per question).

To test the difference of performance in decoding ability using competition as an introduced factor and the game as tool, three groups were created. The Competition Group, the No Competition Group and the Control Group. In the first two groups the game is played using computers. In the Control group, the concept is represented using pen and paper in the form of a simple written questionnaire.

To test the improvement of the skill itself (the gain), it was important to measure the performance of the participants twice. So the study will have two phases. Since we are trying to develop a skill in decoding and inferring information, it was important that the students did not get accustomed to the setup itself. Knowing the answers and setup beforehand defeats the purpose of decoding and inferring. So, it was decided that the test would have different kinds of questions for each of the two phases.

Each battle in the BattleQuiz is composed of seven questions (four answers each, only one being correct), chosen randomly from a set found on the database respective to the virtual arena. For the study, there will be two Arenas, one for

each phase and its respective set of questions and answers.

For each group session in each phase, the game was decided to be played 10 times (70 questions answered). This made the game, brief enough as to not overbear the students, and long enough to have a significant number of playthroughs and data. Having shorter sessions (20 to 30 minutes) also helped in scheduling with the school, in off class hours, between classes, and in incentivising parents to agree to the participation in the study.

Session Setup

The study focuses on the three groups and two phases. In this section we describe the differences between groups and the phase preparations.

Groups

Groups are defined by the game versions they execute in the first phase. There are a total of three groups, the Competition Group, the No Competition Group and the Control Group.

For the first phase the Competition Group has the presence of various stimuli to stimulate a sense of competition between participants while playing the game. The No Competition Group has no such stimuli but still has participants play the game. The Control Group has no competition stimuli and does not play the game, instead, being given a simple written test.

In the second phase all groups play the same version. All of them play with the same conditions of the No Competition Group (no competition stimuli but still playing the game) while maintaining the same participant distribution. Here, the questions and answers are different as well.

Competition Stimuli

To test the presence of competition in the game design setup, some aspects of the App were used. Mainly, the friend system and history data. The application of the setup is explained more practically in the section **test group**.

The friend system in the app, makes it possible for the player to look for and add other existing players of known names. Before the session with the Competition Group, the user IDs were logged in and added to each other with friend requests and, subsequently, accepting the requests.

As the participant's sat down to start the study, they were openly asked for their number (which can be found in the interface profile of the app). Having players know their ID and their friend's ID, makes players perceivable to each other while in the game. All of this coupled with the design component of the friend list, accentuates tension and focus on to each other, adding to the feeling of competition, according to the Social Influence and Accomplishment factors from Octalysis (Yu-Kai Chou: Gamification & Behavioral Design, Octalysis framework) [36].

During the sessions (as described below in the section **test groups**, players would be asked to challenge each other and reply to challenges. The challenging of a class member would add to the competition feeling, but having the other participant reply to the challenge and have a "you win" or "you fail" message show at the end would elevate this feeling greatly, as it would be perceived by their reactions and interactions. Additionally, participants would be encouraged to talk between games, see and compare scores, as interaction is a component of competitions and is a heavy driving factor for learning [23].

This incentivising (stimulating the participants by interacting with them and asking for their scores, difficulty playing) and highlighting (good

scores, and social standing) of social interaction, was all done to appeal to the social influence drive of gamification (Octalysis framework) which is the main influencer of the sense of competition [36].

Now that the test would be different, and students less biased on the second phase, this brought another issue. If the questions are so different, would that not affect their recorded performance, making them not comparable? This is another reason for the existence of the Control Group. To circumvent the issue, this group would have a different *onboarding* (introduction) to the concept of decoding on phase 1, through a written form, instead playing the game only on the phase 2.

In the end of the sessions, participants would be asked to count their wins and defeats, by going to the history section. This was done, because the ranking functionality of the app, sadly, was not working.

Phase Preparation

Phase 1

For the first phase, it was necessary to create many questions, as there should be many students participating and playing 10 times in each. According to the set up, each student will see around 70 questions, possibly some repeated ones, since they are randomly chosen, and it is highly unlikely that no question will be showed twice. Though, 100 would be optimal, only 72 were made (for the first phase), during the time towards the experiment.

These questions were not really questions. The participant would be presented with a sentence (e.g. The girl summed the odd numbers.), and choose an answer that felt most relevant or correct to it (e.g. Even). They were created, focusing on the interpreting and inferring of information from the given sentence and context

of the participant (In this case, the knowledge of a child from Portugal in the city of Porto).

Question and Answer Creation

For questions creation, many things were taken into account for phase 1:

First, it was necessary that the answer to the question was not always of an extremely high *cue validity* (Basic Level categorization, where words are most easily decoded and expected) [8]. When the answer would be very relatable (basic level) to words in the given sentence, the other 3 possible answers would be of equal or higher *cue validity* but not as correct. Examples:

“The anaconda snake embraces the bunny rabbit” - In this situation, it is highly recognizable that the snake is attacking the bunny. So, words such as “fangs”, “speed” and “venom” are also presented, as they are associated with snake attacks. But “choke” is the correct answer.

“The man ran straight, from one point to the other point”. In this case the answer is “segment” (part of a line, connecting two points in space) and the wrong answers would be “race”, “speed” and “legs” which are of a much higher *cue validity* (which are of basic level, comparing to the subordinate correct answer).

Second, the answers could not be too unrecognizable (too superordinate or subordinate) to the participants, in their context (their culture, age (11 year old), students of Porto in Portugal). For this, the sentences were made using the Portuguese Official Program Goals for the different disciplines. This was so that some complexity can be used, while maintaining the basic object validity (which wavers according to the recipient’s knowledge of said area). Example:

“The abuses of those in power leave the people in discontent” - In this case, it would be expected that words such as “society”, “government” or “injustice” would be of basic level and “revolution” (the correct answer) of subordinate level. But for the ages of the participants, “revolution” is one of the big school year subjects

for the discipline of history, as well as being an important national holiday, so it is not unrecognizable

Third, there should be questions of different difficulties, as to give some randomness and unpredictability to the participant’s expectations.

Example:

“The man measured the distance” - The answer is “ruler”, and not “watch”, “thermometer” or “scale”, which is extremely easy and obvious.

Phase 2

For the last phase, it was agreed that only 25 questions should be needed, as, other than taking time, the issue of participants getting accustomed with the format, was mitigated by having a different set of questions and answers, and different format of those questions. This is because this phase will just access the changes and adaptability from the previous phase. In the end 27 questions were created.

The sessions with the participants, would be the same as before, 20 to 30 minutes, going through 10 battles in unison. This was decided as to maintain the schedule flexibility.

Question and Answer Creation

For the 27 question creation of phase 2, there was no need for creation, just reusing and editing materials constructed by Barry S. Stein and John Bransford (1984) [20].

Since this phase’s questions would be focused in the concept of elaboration, questions, in this case, were sentences as well, that needed to be completed. These sentences were taken from the paper “Elaboration and knowledge acquisition”, which tested the use of different kinds of elaboration referenced in the **Context** section [20]. They had 2 appendixes with sentences and respective elaborations according to the precision (how well they justified the first part of the sentence) and relatedness (how the words in the second part of the sentence related to the first part). Since the paper concluded that with the

objective of memory retention, precision was a significant variable, while relatedness was not, all the existing sentences for high precision were used for the study (27 total).

Tests

Dependent Variables

With the setup decided, particular information would be needed for the evaluation of the findings. In this case, the main questions are to evaluate the difference between groups in each phase, in terms of scoring and reaction time.

To calculate these differences, two variables were recorded in the playthroughs of the game. The response time (average for each question replied by each participant) and performance (percentage of correct replies)

Using a one-way ANOVA, we can calculate the significant difference between the performance of the three groups in each of the phases independently.

Variable Recording

When selecting “choose an opponent” to choose an opponent (either another player or the computer, depending on group) and then selecting the arena available (Semantical Education 1 or 2, respective to phases 1 and 2), players start their battle. On each battle a player is presented with 7 questions, 1 at a time. For each questions, they are shown a timer, counting down (30 seconds) that stops and resets when an answer is selected, the response time and importance of the reply are then recorded and sent to the database. Showing the time counting down, is a game design element that gives urgency. The correctness of the answer is based on the predetermined correct answer for each specific question.

Procedures

In the first week and next monday, the three groups had their first phase. In the next week, all groups would have their second phase, which had the same setup and methodology as the **No Competition Group**, maintaining group distribution (and schedule of the first phase).

No Competition Group

Schedule. Phase 1, 29 of May, from 14:00 to 14:25.

Phase 2, 5 of June, from 14:00 to 14:25.

Participants. Expected participants were 9 in total, 4 male and 5 female students, 4 from the 5th grade, 5 from the 6th grade. The turn out was 7 students (2 female students from the 6th grade where absent).

Method. During the lunch break, the session’s supervisor (project’s Master’s student), opened each available computer, opened the browser and opened the page for application page. For each respective computer, he would login with a specific test ID (1 to 30 were available) and record it on the notebook. As students arrived, they would come to the supervisor, to give their names (checking presence) and according to where they sat down, their names would be connected to the respective test ID number (as to maintain anonymity). For their next session, students would be seated in the same computer and test ID.

The supervisor would start by explaining that the game was similar to a tv question gameshow, explaining that in this case the questions would be more arbitrary, and that they just needed to choose the answer that seemed most correct, by logic or feel. He also added that it would be important to not talk between themselves about the questions, as they were different for everyone.

The sessions had the group of students play the game battle 10 times in unison (doing each battle at the same time, waiting for each other to finish and only continuing to the next one when all were finished). They started the first battle (7 questions) against the computer option, and waited for everyone to finish, only then would the second battle start, and so on...

Competition Group

Schedule. Phase 1, 30 of May, from 14:00 to 14:25.

Phase 2, 6 of June, from 14:00 to 14:25.

Participants. Expected participants were 10 in total, 3 male and 7 female students, 6 from the 5th grade, 4 from the 6th grade. The turn out was extremely low, 4 students (3 female and 1 male from the 5th grade). To compensate the small turn out, it was decided that an added session would be performed with the missing students of that day, during the week, 1 of June, from 14:00 to 14:25. The turnout on this day was much greater, and students that were not able to come to other group schedules, joined (sadly, this adds to the unbalance in the group numbers) having a total, of 12 students (9 female, 2 male). If this is shown to be too vast of a discrepancy in number of participants, only the second turnout will be taken into account total 8 (7 female, 1 male).

Method. The overall set up was the same as **No Competition Group**. Adding to it, the following stimuli were introduced during Phase 1:

Before starting, the supervisor instigated a feel of competition, by explaining the existence of a ranking system (high score board). He expanded on how the participants would be playing against each other, and how they would be rewarded points for speed and correctness of their answers. Explaining that he would be curious to see who were the best players (as to point them out).

During the sessions, in this case, instead of challenging the computer, they challenged one of the participants in the room, chosen by the supervisor. While this choice was made, the supervisor would wait a little, letting the students see their next challenge and talk about their scores (instigating a feel of competition). This was done repeatedly, going around the tables, as to get everyone to challenge and respond to the challenges of each other (this was only possible with 10 or less students, as to have everyone challenge each other).

In the end of the session, the participants checked their battle history, counting their wins, to see who had the best performance. One by one, the supervisor would ask for the specific participant's scores and comment on them, while letting them speak freely with each other.

Control Group

Schedule. Phase 1, 4 of June, from 14:00 to 14:25.

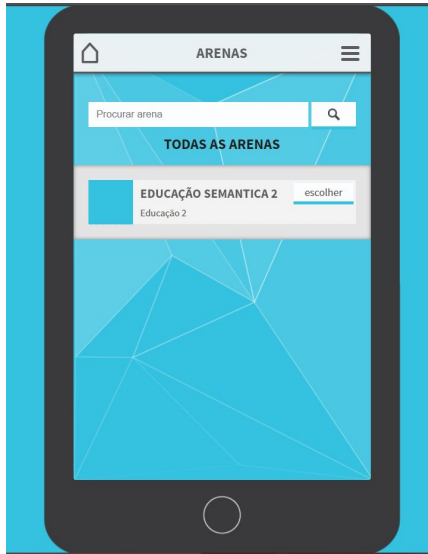
Phase 2, 8 of June, from 14:00 to 14:25.

Participants. Expected participants were 9 in total, 4 male and 5 female students, 3 from the 5th grade, 6 from the 6th grade. The turn out was under the total, with 7 total (4 female, 3 male).

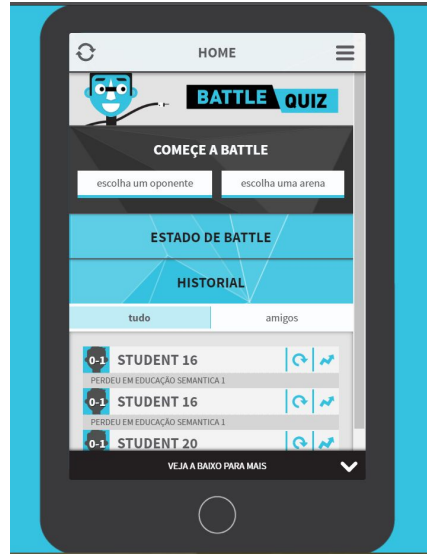
Method. The setup for Phase 2 was the same as the **No Competition Group**. For Phase 1, it was quite different: In this case, the students arrived and sat down in the tables in any order. They were given a sheet of paper with printed questions and answers (written test format) and asked to answer them all (total of 30 random previously chosen questions from phase 1).

The supervisor started by explaining that the these questions were only to know their opinion, and that in the next session they would play the game. He then explained that they just needed to choose the answer that seemed most correct, by logic or feel.

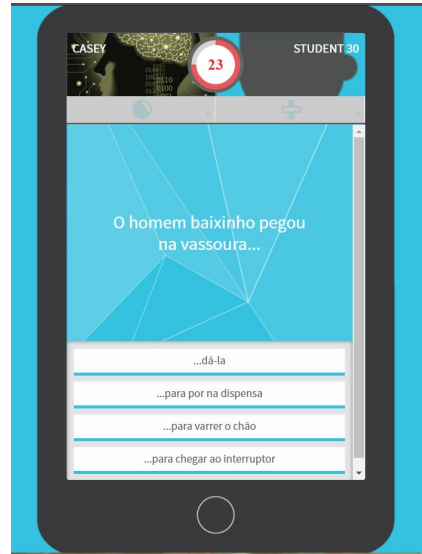
In the end, when gathering the tests, he asked for the name of each student, and appoint an unused test ID to the name and respective test sheet. As to later use the same ID in the application for the second phase (once again, maintaining anonymity).



Arena Selection



Main page, showing history section



Question in battle

Results

The objective amount of 36 students was not accomplished, this will be further explained in the **Discussion**. There were four male students for each of all the groups and five, seven and four female students respective to **No Competition Group**, **Competition Group** and **Control Group**. This lack of balance of female students in the groups was because of some late submissions of more participants.

Statistics were received from Liftov in the shape of excel table files. Each table had multiple columns but only the relevant ones will be taken into consideration and presented in this document.

Each line was composed of the ID of the participant and the ID of the question answered. For each question, there was the amount of times it was presented, followed by the amount of times it was correctly answered and then respectively calculated percentage. IT also had the average time of response and the Phase ID.

To proceed with the data analyses, for each student, the average correct percentage and response time of all questions was calculated, while still separating phases.

	Mean	Std. Deviation	Std. Error
No Competition	71.6057	17.30246	6.53971
Competition	80.6513	8.03162	2.83960
Control	70.1083	15.00671	6.12646
Total	74.6238	13.87542	3.02786

Table 5: Overall scores for Phase 2 (one-way ANOVA)

For the second phase, a one-way ANOVA was used to compare the mean scores. There was not a significant effect of competition on scores at the $p < .05$ level for the three conditions [$F(2,18) = 1.272, p = 0.304$]. Because there is no significance in the results, there was no need to compute the Post Hoc tests.

	Mean	Std. Deviation	Std. Error
No Competition	2.15938	0.778574	0.294273
Competition	2.12608	0.413145	0.146069
Control	2.79630	0.521452	0.212882
Total	2.32867	0.633736	0.138293

Table 6: Overall reaction time for Phase 2 (one-way ANOVA)

To compare the reaction time in Phase 2, a one-way ANOVA was used, as this time, the **Control Group** played the same game and could record the reaction time. There was not a significant effect of competition on scores at the $p < .05$ level for the three conditions [$F(2,18) = 2.676, p = 0.096$]. Because there is no significance in the results, there was no need to compute the Post Hoc tests.

Summary

It can be concluded with these values, that the results of the hypothesis for the performance of the different groups was inconclusive.

None of the groups showed to be comparable in phase 2, thus reliable comparisons cannot be made yet. These inconsistencies will be further expanded in the Discussion area.

Discussion

Some of the hurdles found during the development and preparation for the study tests were:

- Too many playthroughs of the game with the current amount of questions can be rend the data in each phase untrustworthy. This is because, unexpectedly, the players quickly started to memorize the answers to the questions (according to the feedback). They would act using memory and not interpretation starting from the 5th or later game, according to their testimony. It is important to give some kind of feedback to the players so they can learn, so removing the feedback (“correct” or “incorrect” screen response) would not help. But increasing the amount of questions and decreasing the amount of playthroughs should be a good possible workaround.
- The question and sentence creation for each phase is quite lengthy and cumbersome and should not be made alone, like it was in this study. It is imperative for the workload and for the quality assurance to have more hands in the creation, mitigating partiality. 200 questions should work fine for a future study (100 for each phase).
- The small group turn out for participants comes from a bad academic year timing and understanding of it. For a future study, it will be very important to contact and work with a school at the start of the year. Preparing, with their help, to perform the study more than two months after, which was the case of this study as the academic year was about to finish. Consider for the scheduling of the study, the tests, exams and other evaluation periods of the target grades. Visit the school’s resources (classrooms and computer rooms) and schedule talks to raise awareness of your study to the students and to the parents. It is

imperative to connect to students and show them how this could change techniques of education, using games, as they will be the primary force to have parents accept their participation. More than one visit will help students remember and make sure all permissions are given out. Teacher awareness will also help with the student receiving of submissions.

- The study performance itself should have either someone from the school or more than one study representative in the room. Even though someone from the school would better help deal with the students and to incentivise the right mood when necessary.

How the sentences should be created to analyse semantic decoding in the future

In future experiments, a different style should be adopted for testing in the first phase. Looking to enhance the retrieval of relevant features in the working memory, it will be important to connect to concepts in the long term memory.

Words that, in context, point to less frequent semantic interpretations or meanings are harder to retrieve and connect. So, it is important to activate them and have the reader make the connection, acknowledging it. This does not mean that connecting to the most frequent or intuitive meanings of a word is not important.

As mentioned before in this study, it is important to have in mind that the semantic value of a word can not only change according to the surrounding sentence but also to the context around the reader. So, new sentences for the game should be made as they will probably be used in the Netherlands and not in Portugal, as the culture and common knowledge for a child growing up in different countries will be

different, even in this day and age of world interconnectivity and fashions.

In phase 1, each question (example sentence) should be followed by the connection between one of the words in it to a semantic meaning (preferably one word). Asking the player to categorize its importance to the sentence.

E.g. The sentence “The man pushed the piano to the next room.” would be followed by the connections “Piano - Heavy”. In this case the correct answer would be that this connection is indeed “Important”, even if Piano is more easily connected to instrument, in this context that is not the most relevant feature of Piano.

The second phase should stay the same as the one performed in this study. By having players classify a proposed sentence continuation, they will be choosing the context itself.

Again, it is important, in the sentence creation, to have outliers, to challenge the player’s thinking and rationale. Which, in the first phase would activate or connect a word to its most typical and intuitive meanings. And in the second phase, have wrong and illogical proposals for continuations.

Group Expectations

The control group should be well familiarized with the kind of questions and answers, but without the game element. It is expected, despite the lack of significant finding in this study, that the control group will be slower to find the right answer, compared to the game groups.

Competition drive higher engagement, creating attention to this game’s stimuli. So it is also expected that future work will reveal that the competitive group will have faster and better results than the non competitive group. The tested groups were very smalls, and it is speculated that that is why it was difficult to find

significant data. But with the information given in this study about the logistics of the experiment, it will be easier to have a bigger and more concise study group.

Game Design elements in the developed app

Visually, the questionnaire type game, presenting buttons to press is an intuitive game element for children today.

The Challenge list uses the motivation concept of competition, giving a name and face to the challenger, appealing to the social and connected environment.

The History list appeals to the social hierarchy motivation, as it gives the player a interpretable performance in facing opponents.

In each game and question, the points given help give feedback to the speed of answering and if it correct. Motivating the player to try and be faster and better, as there will be a clear change in the number of points awarded in each reply.

Their interactions with each other and the supervisor showed their engagement and motivation when learning with the app (both the video-game rules and other knowledge).

The study's research questions and hypothesis stands.

Conclusions

Overall results did not have enough significance to confirm or deny expectations, but served as a pilot study for the recording of all the necessary steps to ensure a more successful study in the future. The competition factor was not significant enough for a conclusion. The "video game" component's connection to learning precise elaboration also did not show conclusive evidence.

With a bigger pool of students and different but still context appropriate sentences, significant evidence is expected.

Still, the game, competition components and questions used, showed to grab the students attention and interest in a very positive way.

Reference

- [1] John D. Bransford, "Human Cognition: Learning Understanding, and Remembering", 1979
- [2] Prof. Fernanda L. Viana and Maria M. Teixeira, "Aprender a ler, da aprendizagem informal à aprendizagem formal", 1st edition, 2002
- [3] Emerald V. Dechant, Henry P. Smith, "Psychology in teaching reading", 1977
- [4] Lobrot M., "Lire", 1980
- [5] Inês Sim-Sim, Cláudia Ramos, Eveline Silva, Manuela Micaelo, Maria M. Santos and Palmira Rodrigues, "Ler e ensinar a Ler", 1st edition, March 2006
- [6] Lesgold et Al., "Computer resources for learning", 1985
- [7] LaBerge D. & Samuels, 3rd edition, S. J., "Toward a theory of automatic information processing in reading. In Theoretical Models and Processes of Reading", 1985
- [8] Eleanor Rosch, Carolyn B. Mervis, Wayne D. Gray, David M. Johnson, and Penny Boyes-Braem, "Basic Objects in Natural Categories", 1976
- [9] Morris, Bransford and Franks, "Levels of processing versus transfer appropriate processing", 1977
- [10] Gordon H. Bower, Michal C. Clark, Alan M. Lesgold, David Winzenz, "Hierarchical retrieval schemes in recall of categorized word lists", 1969
- [11] Tulving E. and Shirley Osler, "Effectiveness of Retrieval Cues in MEemory for Words", 1968
- [12] Tulving E., Thomson D. M., "Encoding specificity and retrieval processes in episodic memory", 1973
- [13] Tulving E. and Wiseman S., "Relation between recognition and recognition failure of recallable words", 1975
- [14] Watkins M.J., & Tulving E., "When retrieval cueing fails", 1978
- [15] Flexser AJ, Tulving E., "Retrieval independence in recognition and recall Psychological Review", 1978
- [16] Allan Paivio, "Imagery and deep structure in the recall of English nominalizations Journal of Verbal Learning and Verbal Behavior", 1971
- [17] Allan Paivio, "Imagery and synchronic thinking", 1975
- [18] Allan Paivio, "Concerning dual-coding and simultaneous-successive processing", 1976
- [19] Barry S. Stein and John D. Bransford, "Constraints on Effective Elaboration", 1979
- [20] Barry S. Stein and John Bransford, "Elaboration and Knowledge Acquisition", 1984
- [21] Barry S. Stein, Karla F. Brock, Donny R. Ballard, "Constraints on effective pictorial and verbal elaboration", 1987
- [22] Ermi & Mäyrä, "Measuring and defining the experience of immersion in games", 2005
- [23] Timothy K. O'Mahony, Nancy J. Vye, John D. Bransford, and Elizabeth A. Sanders, "A Comparison of lecture-based and challenge-based learning in a workplace setting", 2011
- [24] Paul Cairn, Anna L. Cox, Matthew Day, Hayley Martin, Thomas Perryman, "Who but not where: The effect of social play on immersion in digital games", 2013
- [25] Norman A. Sprinthall, Richard C. Sprinthall, "Educational Psychology: A Development Approach", 6th Edition, 1993
- [26] Bruner J. S., "Towards a Theory of Instruction", 1966
- [27] Deutch M., "The effects of cooperation and competition upon group process", 1949
- [28] White R., "Motivation reconsidered: The concept of competence", 1959

- [29] Minas J. S., Scodel A., Marlowe D. and Rawson H., “Some descriptive aspects of two-person non-zero-sum games”, 1960
- [30] Judith ter Vrugte, Ton de Jong, Sylke Vandercruysse, Pieter Wouters, Herre van Oostendorp, Jan Elen, “How competition and heterogeneous collaboration interact in prevocational game-based mathematics education”, 2015
- [31] George R. Potts, “Information processing strategies used in the encoding of linear orderings”, 1972
- [32] George R. Potts, Karl W. Scholz, “The internal representation of a three-term series problem”, 1974
- [33] Pollchik A., “The use of embedded questions in the facilitation of productive learning”, unpublished dissertation, 1975
- [34] George R. Klare, “The measurement of readability”, 1963
- [35] Edward L. Thorndike and Irving Lorge, “The teacher’s word book of 30,000 word”, 1944
- [36] Yu-Kai Chou, “Gamification & Behavioral Design, Octalysis framework”, 2018, available at <https://yukaichou.com/gamification-examples/octalysis-complete-gamification-framework/>

Tables

Multiple Comparisons

Dependent Variable: Overall Score for Phase 1

Tukey HSD

(I) Group	(J) Group	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
No Competition	Competition	-.81339	6.33396	.991	-16.9787	15.3519
	Control	-15.49214	6.80880	.085	-32.8693	1.8850
Competition	No Competition	.81339	6.33396	.991	-15.3519	16.9787
	Control	-14.67875	6.60947	.095	-31.5472	2.1897
Control	No Competition	15.49214	6.80880	.085	-1.8850	32.8693
	Competition	14.67875	6.60947	.095	-2.1897	31.5472

Multiple Comparisons

Dependent Variable: Overall Score for Phase 2

Tukey HSD

(I) Group	(J) Group	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
No Competition	Competition	-9.04554	7.08558	.426	-27.1291	9.0380
	Control	1.49738	7.61677	.979	-17.9419	20.9366
Competition	No Competition	9.04554	7.08558	.426	-9.0380	27.1291
	Control	10.54292	7.39379	.349	-8.3272	29.4131
Control	No Competition	-1.49738	7.61677	.979	-20.9366	17.9419
	Competition	-10.54292	7.39379	.349	-29.4131	8.3272

Multiple Comparisons

Dependent Variable: Overall Reaction time for Phase 2

Tukey HSD

(I) Group	(J) Group	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
No Competition	Competition	.033300	.303537	.993	-.74138	.80798
	Control	-.636914	.326292	.153	-1.46967	.19584
Competition	No Competition	-.033300	.303537	.993	-.80798	.74138
	Control	-.670214	.316740	.115	-1.47859	.13816
Control	No Competition	.636914	.326292	.153	-.19584	1.46967
	Competition	.670214	.316740	.115	-.13816	1.47859

Appendix

Appendix 1 - Phase 1 Questions

In this section you will find the translation (portuguese to english) of the questions created for the study. “Q” representing the sentence presented to the participant, “R” representing the correct response and “W” the wrong responses.

Q:The sheep herder always has to goes downhill when going back to the village.

R: Valley
W: Bay
W: Plain
W: Forest

Q:A penguin sunbathes in antarctica

R: Cold
W: Warm
W: Bird
W: Aquatic

Q: Que Queen drinks fresh tea from Azores.

R: Warm
W: Meal
W: Country
W: Leader

Q: The bus stops in Marques de Pombal.

R: Town square in Porto
W: Lisbon
W: Earthquake
W: City

Q: The children were taught about the importance of Marques de Pombal

R: Earthquake
W: Town square in Porto
W: Bird
W: City

Q:Many foreigners move to Algarve for their retirement

R: Immigration (immigrating to this country)
W: Emmigration (immigration to another country)
W: Tropical
W: Vacation

Q:Many nurses move to the United Kingdom for work.

R: Emmigration (immigrating to another country)
W: Immigration (immigration to this country)
W: Culture
W: Vacation

Q:The seniors and the children look for something that connects them.

R: Adulthood.
W: Death
W: Infancy
W: Birth

Q:One child is born in 1500, another is born today.

R: Longevity
W: Crying
W: Milk
W: Caring

Q:A man makes a phone call to China.

R: Instantaneous
W: Distance
W: Business
W: Discussion

Q:The statue of D.Afonso Henriques is positioned in the middle of the citysquare.

R: Patrimony
W: War
W: Weight
W: Sculpting

Q:A statue is built to the exact likeness of the old king.

R: Sculpting
W: War
W: Patrimony
W: Weight

Q:The men, push the car, out of charge into the garage.

R: Weight
W: Transportation
W: Race
W: Trip

Q:The hunter hunts illegally in an endangered area

R: Extincion

W: Sport
W: Elk
W: Food

Q: The girl summed the odd numbers
R: Even
W: Odd
W: Negative
W: Addition

Q: The sun is at its highest at midday.
R: Perpendicular
W: Morning
W: Obtuse
W: Afternoon

Q: The sprinter runs straight, from one point to another point.
R: Segment
W: Race
W: Speed
W: Legs

Q: The eskimo tries to drink water
R: Ice
W: Vapor
W: Swim
W: Fishing

Q: The paleontologist found a dinosaur jaw, it doesn't have fangs.
R: Herbivore
W: Carnivorous
W: Hunter
W: History

Q: The paleontologist found a dinosaur jaw, its fangs are quite big.
R: Carnivorous
W: Herbivore
W: Hunter
W: History

Q: The Iberian wolf is becoming more and more rare
R: Extinction
W: Canine
W: Pack
W: Peninsula

Q: Comes down from the ceiling but doesn't fall.
R: Stalactite
W: Rain
W: Stalagmite
W: Leaf

Q: Traveling from horizon to horizon, from East to West, keeping distance with the Poles.
R: Equator
W: Magnetism
W: Hemisphere
W: Wind-rose

Q: Breathes but does not move
R: Plant
W: Mineral
W: Fish
W: Mammal

Q: The biggest animal that eats the smallest.
R: Blue whale
W: Plancton
W: Lion
W: Elephant

Q: The fly-trap plant waits for an insect to enter its leaves to capture and digest it.
R: Trap
W: Search
W: Hunt
W: Mobility

Q: The butterfly shows the big eye pattern in its wings to fool predators.
R: Intimidation
W: Hiding
W: Camouflage
W: Escape

Q: The peacock shows its colourful feathers to impress females.
R: Reproduction
W: Hiding
W: Intimidation
W: Camouflage

Q: The old turtle travels from one side to the other of the island.
R: Slow
W: Hiding
W: Speed
W: Defense

Q: The anaconda snake embraces the rabbit.
R: Strangling
W: Venom
W: Fangs
W: Speed

Q:The sonar is used to find prey in the darkness.

R: Bat

W: Dolphin

W: Submarine

W: Fly

Q:The sonar is used to map the deepsea floor and enemies.

R: Submarine

W: Dolphin

W: Bat

W: Cargo ship

Q:The caterpillar form a cocoon around itself.

R: Metamorphosis

W: Evolution

W: Hibernation

W: Protection

Q:Penguins always keep their eggs between their legs

R: Hatching

W: Mobility

W: Food

W: Defense

Q:The octopus, feeling threatened, lets out a cloud of ink.

R: Escape

W: Defense

W: Reproduction

W: Attack

Q:The twins were born, one girl and one boy

R: False twins

W: Triplets

W: True twins

W: Half-Brothers

Q:The titanic hit an iceberg

R: Historical disaster

W: Trip

W: Ship

W: Vacation

Q:The navigator, sad for not being able to simply cross the small area of land with his boat, will have to go around by sea.

R: Peninsula

W: Taxes

W: Discovery

W: Commerce

Q:The race in the boulevard starts, the competitors are an even match, staying side by side during the race.

R: Paralel

W: Circuit

W: Perpendicular

W: Overtake

Q:A small pebble is put in a water bucket, but the water doesn't rise as much and when the bigger rock was put before.

R: Volume

W: Weight

W: Floating

W: Shape

Q:The man was taken from his home, to work without pay, against his will.

R: Slavery

W: Culture

W: Job

W:Rights.

Q:The abuses of the people in power leave the common people unhappy.

R: Revolution

W: Society

W: Democracy

W: Patience

Q:The woman doesn't like dogs, every time they are near she sneezes.

R: Alergie

W: Intolerance

W: Cats

W: Fear

Q:The man is losing weight because of his balanced meals.

R: Nutrition wheel

W: Exercise

W: Cirugy

W: Hungry

Q:The boy feels awkward because pimples have started appearing.

R: Puberty

W: Itchy

W: Ugly

W: Adult

Q:The elephant scared the dog.

R: Big

W: Small

W: Violent

W: Fat

Q:A big flea jumped onto the back of the dog, but it didn't even notice.

R:Small

W: Big
W: Calm
W: Violent

Q:The youtube channel got recognized by bad publicity
R: Drama
W: Scandal
W: Youtuber
W: Lack of commercials

Q:The youtube video was controversial and had a lot of views, still it might not have been successful.
R: Lack of commercials
W: Scandal
W: Drama
W: Youtuber

Q:The man measured the weight.
R: Weighter
W: Clock
W: Ruler
W: Thermometer

Q:The man measured the time.
R: Clock
W: Weighter
W: Ruler
W: Thermometer

Q:The man measured the distance.
R: Ruler
W: Clock
W: Ruler
W: Weighter

Q:The man measured the temperature.
R: Thermometer
W: Clock
W: Weighter
W: Ruler

Q:The photo has a small amount of pixels
R: Bad definition
W: Digital
W: High-resolution
W: Old

Q:The constant use of fossil fuels, will lead to scarcity.
R: Non renewable energy sources.
W: Non pollutant
W: Renewable
W: Atmosphere

Q:Even though there isn't always wind, windmills, will always catch it.
R: Renewable energies
W: Non pollutant
W: Non renewable
W: Atmosphere

Q:The wind come naturally, so it does not leave any substances.
R: Non pollutant
W: Non renewable
W: Atmosphere
W: Pollutant

Q:Batteries have toxic chemicals inside. Careful when throwing away.
R: Polluting
W: General trash
W: Atmosphere
W: Non renewable

Q:The narrator presents all events, including the thoughts of the characters.
R: Omniscient
W: Literal
W: Subjective
W: Objective

Q:The man used a measuring tape in his job.
R: Tailor
W: Infrastructure
W: Math
W: Drawing

Q:The man went to the airport by taxi.
R: Departure
W: Flight
W: Arrival
W: Trip

English language studies

Q:The english man colored his hair blonde
R: Blonde
W: Yellow
W: Ginger
W: Brown

Q:The english man lifts the heaviest weight at the gym
R: Strong
W: Sports
W: Weak
W: Quick

Q:The english man waits at the bus stop

R: But stop
W: Taxi
W: Carpool
W: Transportation

Q: The english man has blue eyes
R: Blue
W: Blond hair
W: Blue hair
W: Brown eyebrows

Q: The english man hits his head on the top of the doorway.
R: Tall
W: Thin
W: Fat
W: Short

Q: The english man can't reach the shelf.
R: Short
W: Thin
W: Tall
W: Fat

Q: The english man points his finger at the other
R: Rude
W: Finger

W: Secret
W: Hand

Q: The english woman has curly hair
R: Curly hair
W: Straight hair
W: Wavy hair
W: Short hair

Q: The english woman lost too much weight
R: Thin
W: Short
W: Tall
W: Fat

Q: The english man runs slowly and quickly gets tired
R: Fat
W: Tall
W: Short
W: Thin

Q: The english boy is thankful for the present
R: Thankful
W: Child
W: Short
W: Adult

Appendix 2 - Phase 2 Questions

Q: The short man lifted the tent...

R: ...which was half a meter tall

W: ...in the forest

W: ...in the circus

W: ...quickly

Q: The short man picked up the broom...

R: ...to reach the light switch

W: ...to sweep the floor

W: ...to put in the storage

W: ...to give it

Q: The funny man bought a bow...

R: ...which shoots water

W: ...to wear with his suit

W: ...for his collection

W: ...to give it as a present

Q: The funny man bought a tube...

R: ...to make soup for the party

W: ...to take a bath

W: ...to clean his dog

W: ...to be able to lie down when bathing

Q: The strong man read the book...

R: ...about how to use weights at the gym

W: ...about a romance

W: ...about magic

W: ...about chemistry

Q: The man with white hair has a flask...

R: ...of hair dye

W: ...of poison

W: ...of juice for the trip

W: ...of nail polish

Q: The fat man went to the train station...

R: ...and got stuck in the entrance gates

W: ...but the train didn't come

W: ...but his friends were late

W: ...and arrived just in time

Q: The fat man wore his shoes in the street...

R: ...and they got worn out in a month

W: ...and got dirty

W: ...and the colors fit with his clothing

W: ...but the shoelaces were untied

Q: The rich man bought the paint...

R: ...change the color of his luxurious boat

W: ...to put in his printer

W: ...to paint the wall in his home

W: ...to save for the constructions later

Q: The rich man went to get the chair...

R: ...to use as lumber for the fireplace

W: ...for his guests to sit on

W: ...to eat in the kitchen

W: ...to reach the top shelf

Q: The hungry man bought the tie...

R: ...to get into the luxurious restaurant

W: ...for the business meeting

W: ...to go to court

W: ...to go to the party

Q: The tall man bought the cookies...

R: ...that were in the top shelf

W: ...chocolate flavored

W: ...that were the cheapest

W: ...that were the most expensive

Q: The tall man went to the barbershop for a haircut...

R: ...the barber had to get on top of a stool

W: ...and then helped the barber clean up afterwards

W: ...but didn't like his new haircut

W: ...but didn't like the price of the haircut

Q: The smart man went to work...

R: ...before the traffic hour

W: ...but forgot his house keys

W: ...by tram

W: ...and arrived late

Q: The smart man used the coal...

R: ...when the pencil broke

W: ...to cook dinner

W: ...in the fireplace

W: ...as a paperweight

Q: The bald man used his phone ...

R: ...to check sales on hats online

W: ...to call his mother

W: ...to go to facebook

W: ...to look at his emails

Q: The sad man admired his new boat...

R: ...but now has to fix the big hole on the boat's side

W: ...before going to a sea trip

W: ...before going fishing

W: ...while cleaning it

Q: The lucky man walked home...
R: ...right before starting to rain
W: ...but he tripped on a rock
W: ...but arrived earlier than he expected
W: ...and got lost

Q: The short man bought a broom...
R: ...to sweep the floor in the attic
W: ...to sweep the kitchen floor
W: ...to have in the storage
W: ...to sweep the small crumbs

Q: The brave man gave the money...
R: ...before the burglar shot him
Q: ...to save a sick person
Q: ...because he was scared of the curse
Q: ...to the church offering

Q: The fat man read the sign...
R: ..."careful with thin ice"
Q: ..."don't step on the grass"
Q: ..."careful with the dog"
Q: ..."no pictures"

Q: The thin man found the scissors...
R: ...to cut his belt in half
Q: ...to cut the newspaper discounts
Q: ...to open the cookie package
Q: ...to cut the cleaning instructions in the new clothing

Q: The rich man picked up the chair...
R: ...to see if its legs were golden
Q: ...to put in the empty room

Q: ...as to do exercise
Q: ...to clean the floor

Q: The sad man looked at his new car...
R: ...that got destroyed on the fire
Q: ...proud of its beauty
Q: ...thinking where his next trip will take him
Q: ...thinking of the nice discount he got on it

Q: The friendly man ate dinner...
R: ...and offered to help with cleaning the dishes
Q: ...and sat on the couch, talking
Q: ...while speaking on the phone
Q: ...and left without paying

Q: The smart man went to work ...
R: ...and made a good investment
Q: ...but arrived late for a meeting
Q: ...to the productive factory
Q: ...extra hours

Q: The bald man used the house phone...
R: ...to ask for the discounts on hats
Q: ...to call the plumber
Q: ...to contact some friends
Q: ...to order a pizza

Appendix 3 - Future sentence example

Q: Que Queen drinks fresh tea. Fresh - Warm:

R: Important

W: Not Important

Q:The man moves the piano to the next room. Piano -

Heavy:

R: Important

W: Not Important

Q:The seniors and the children look for something that connects them. Children - Age:

R: Important

W: Not Important

Q:A child is born in 1500, another is born today. Today - Longevity:

R: Not Important

W: Important

Q:A man makes a phone call to China. China - Instantaneous:

R: Important

W: Not Important

Q:The man was taken from his home, to work without pay, against his will. Will - Slavery:

R: Important

W: Not Important

Q:The abuses of the people in power leave the common people unhappy. Abuse - Revolution:

R: Important

W: Not Important

Q:The woman doesn't like dogs, every time they are near she sneezes. Sneezes - Allergy:

R: Important

W: Not Important

Q:The man is losing weight because of his balanced meals. Meals - Exercise:

R: Not Important

W: Important

Q:The boy feels awkward because pimples have started appearing. Pimples - Adult:

R: Not Important

W: Important

Q:The elephant scared the dog. Elephant - Big:

R: Important

W: Not Important

Q:A big flea jumped onto the back of the dog, but it didn't even notice it. Dog - small:

R: Not Important

W: Important

Appendix 4 - Developed Game App

Here the setup for the new game app developed after the pilot study is described. Starting by announcing the **IDE's** (Integrated Development Environment) used, followed by the **Runtime Environments** and afterwards the **Code architecture files**, ending with how the game **Start Up** (how it should be run).

IDE's

Brackets: Used to develop the HTML and javascript used in the **frontend** development. This tool was used because of its ease of code deployment.

VSCODE: Used to develop the javascript for the **backend** of web application.

MYSQL Workbench: Used for creating and accessing a MYSQL Database table and query development and testing.

RunTime Environments

NODE.js: Environment capable of executing the **backend** code using javascript.

NPM: Package manager for Javascript, must be used with Node.js.

MYSQL: Relational Database Manager System, used for storing the application data. Should be installed in the system used for the database server.

Code architecture files

Attached to this thesis, will be a zip file, carrying the used code for the following application sections:

Front end: HTML, Javascript and Css files.

Backend: Web API developed in Javascript.

DDL (Data Definition Language): SQL file with all the commands necessary for creating and initiating the database.

DML (Data Manipulation Language): SQL file with all the Insert queries used to add content to the application.

Using the MYSQL Workbench, one can select the whole file and run it to create and initially populate the database.

Start Up

Node.js: In VSCODE, start by accessing the terminal and running the command “node server.js” which will compile the code and execute. Note that all the backend files from the folder should be present.

MySQL: Make sure that the previously setup database is running. This can be checked in the Services table for Windows.

Browser running and testing:

Brackets runs a continuously live updated browser page that generates a random port at the start of its runtime. This port should be updated in the Backend file when testing.

If the port is not updated, the following error message “CORS policy” will occur with the warning “'Access-Control-Allow-Origin' header has a value 'http://127.0.0.1' that is not equal to the supplied origin”. Other outputs from the Backend will alert and inform of backend issues or errors.

