FACULDADE DE ENGENHARIA DA UNIVERSIDADE DO PORTO

Multimodal Memory Games: Impact of Visual, Auditory, and Haptic Stimuli on Cognitive Load and Player Performance

Martim Miranda Maciel

U.PORTO
FEUP FACULDADE DE ENGENHARIA UNIVERSIDADE DO PORTO

Mestrado em Engenharia Informática e Computação

Supervisor: Prof. Rui Pedro Amaral Rodrigues

Second Supervisor: Prof. Gilberto Bernardes de Almeida

July 29, 2026

Multimodal Memory Games: Impact of Visual, Auditory, and Haptic Stimuli on Cognitive Load and Player Performance

Martim Miranda Maciel

Mestrado em Engenharia Informática e Computação

Approved in oral examination by the committee:

President: Prof. António Coelho
Referee: Prof. Tânia Rocha

July 29, 2026

Resumo

Esta dissertação estudou como diferentes tipos de estímulos (visuais, auditivos e hápticos) afetam a carga cognitiva e o desempenho do jogador em jogos de memória multimodais, com os estímulos a serem usados tanto individualmente como em grupos. Apesar do aumento dos estudos relativos a estas interações, existe uma falha na investigação sobre como estes três estímulos funcionam em conjunto, quer seja em grupos de dois ou todos os três combinados.

Para corrigir esta falha, foi desenhado e desenvolvido um jogo de memória multimodal que integra estes estímulos em várias combinações. Este jogo contêm dois tipos de tarefas relacionadas com a memória: uma tarefa de memorização de sequências (similar ao jogo *Simon Says*) e uma tarefa clássica de fazer pares de cartas. Foi realizado um estudo com 36 participantes, onde cada utilizador completou um conjunto de níveis por tarefa, com cada nível a usar um conjunto diferente de estímulos. A carga cognitiva foi avaliada através de métricas de rastreio visual e do questionário NASA-TLX. O desempenho do jogador foi obtido através do número de erros e do tempo de conclusão de cada nível.

Os resultados indicaram que, quando se usava apenas um estímulo, a condição visual resultou numa menor carga cognitiva e no melhor desempenho, enquanto a condição háptica teve a maior carga cognitiva e o pior desempenho, com estas diferenças a serem estatisticamente significativas, confirmadas através de análises post-hoc de Wilcoxon. O estímulo auditivo apresentou resultados moderados entre as duas condições. Embora as condições de estímulos combinados tenham mostrado melhores resultados em várias métricas, a análise post-hoc não revelou diferenças estatisticamente significativas entre o estímulo visual sozinho e a condição com todos os estímulos combinados, sugerindo que o estímulo visual é o fator dominante e que a combinação de estímulos não melhora significativamente os resultados obtidos com o estímulo visual sozinho, embora não prejudique o desempenho. Estes resultados sugerem que modelos multimodais devem ser abordados com cuidado, de modo a garantir que a inclusão de estímulos adicionais seja justificada pelo contexto de cada tarefa, em vez de ser imediatamente assumida como benéfica.

Abstract

This dissertation investigates how different types of stimuli (visual, auditory, and haptic) affect cognitive load and player performance in multimodal memory games, with the stimuli being used either alone or combined. Despite the growing study of these interactions, there is a lack of comprehensive research on how these three stimuli work together, be it in groups of two or all three combined.

To address this gap, we designed and developed a multimodal memory game that incorporates these stimuli in various combinations. This game contained two types of memory related tasks: a sequence recall task (similar to the *Simon Says* game) and a classic memory card matching task. A user study with 36 participants was conducted, where each user completed a set number of levels per task, with each level using a different set of stimuli. Cognitive load was assessed through eye-tracking metrics and the NASA-TLX questionnaire. Player performance was measured by error count and task completion time.

Results indicated that when using only one stimulus, the visual condition resulted in the lowest cognitive load and best performance, while the haptic condition led to the highest cognitive load and worst performance, with statistically significant differences confirmed via Wilcoxon post-hoc analyses. Auditory stimulus showed moderate results between the two. While combined stimulus conditions showed better results in several metrics, post-hoc analysis revealed no statistically significant difference between the visual stimulus alone and the full combined condition, suggesting that visual stimulus is the dominant factor and that combining stimuli does not significantly improve upon visual alone, though it does not hinder performance either. These results suggest that multimodal designs should be approached with care, ensuring that the inclusion of additional stimuli is justified by the specific task context rather than immediately assumed to be beneficial.

UN Sustainable Development Goals

The United Nations Sustainable Development Goals (SDGs) provide a global framework for achieving a better and more sustainable future for all. It includes 17 goals[1] to address the world's most pressing challenges, including poverty, inequality, climate change, environmental degradation, peace, and justice.

The specific Sustainable Development Goals below are those that obtained the highest support by the contributions of this project.

SDG 4 Ensure inclusive and equitable quality education and promote lifelong learning opportunities for all

SDG 9 Build resilient infrastructure, promote inclusive and sustainable industrialization and foster innovation

SGD	Target	Contribution	Performance Indicators and Metrics
4	4.4	Provides evidence on how external stimuli influence cognitive load and player performance in memory related tasks, supporting the use of these stimuli in the development of serious games for education.	Improvement of task completion time, reduction of error count and lower cognitive load across multimodal scenarios
9	9.5	Advances research in human computer interaction (HCI) through the design and evaluation of multimodal memory games integrated with visual, auditory, and haptic stimuli.	Development and evaluation of multimodal games and the comparison of seven stimuli conditions for cognitive load and performance analysis
	9.b	Promotes innovation in multimodal interaction systems by demonstrating the positive and negative impact external stimuli can have on cognitive load and performance	Integration of three stimuli, alone and combined. These were then implemented on two different tasks to evaluate their effectiveness

[1] https://sdgs.un.org/goals

Acknowledgements

I would like to thank my family and friends for their support and encouragement throughout this work. They kept me motivated and inspired, even when I faced challenges and obstacles, and were always there to help me out when I needed it. To my friends, some of which were also busy with their dissertations, who still found time to help and support my work. A special thanks to Alexandre and Pedro, who did this master's journey alongside me, making it more enjoyable and fun.

To my supervisors, Prof. Rui Rodrigues and Prof. Gilberto Almeida, whose guidance helped me steer myself in the right direction, and whose feedback and advice were always valuable in improving the quality of my work. They were always available to assist me and provide the needed equipment for the development and evaluation of my work. I am grateful for their patience and understanding, and for the opportunity given to learn more.

Martim Miranda Maciel

Declaration of Honour

I declare that this dissertation is my original work and has not previously been submitted or assessed in any other course or curricular unit, whether at this or any other institution.
References to other authors (statements, ideas, thoughts), as well as the use of published works fully complies with the rules of attribution and are appropriately cited in the text and bibliography, in accordance with reference standards. I am aware that plagiarism and self-plagiarism constitute academic misconduct.
Artificial Intelligence (AI) tools have been used in this dissertation to assist in the writing of this dissertation. These AI tools were mainly used to review the text for grammar and spelling errors, as well as to improve clarity and coherence in the writing. All AI generated content was reviewed and analyzed for relevance, accuracy, and appropriateness before being incorporated into the final manuscript, to maintain the scholarly integrity of the research.
The author takes full responsibility for the content of the document.

Martim Miranda Maciel

Martim Miranda Maciel

Table of Contents

List of Figures

List of Tables

List of Acronyms

JSON	JavaScript Object Notation
HCI	Human Computer Interaction
TCT	Task Completion Time
EEG	Electroencephalography
CLT	Cognitive Load Theory
AR	Augmented Reality
RQ	Research Question
VR	Virtual Reality
SG	Serious Games
NASA-TLX	NASA Task Load Index
JSON	JavaScript Object Notation
M	Mean
SD	Standard Deviation
SS	Simon Says
MG	Memory Game
DV	Dependent Variable

Chapter 1

Introduction

There are a wide variety of games designed with different levels of interactions. There are cases where only two senses are used, which is the case of bimodal videogames. In these, the videogame relies on two senses to transmit information to the player, such as visual or auditory cues. Examples of these games date back to the early days, where videogames such as *Pong* relied only on their visuals and sounds to communicate with the player.

Different from these bimodal videogames, multimodal videogames require the use of multiple stimuli transmitted through the videogame to the player to create immersive and engaging experiences. These can range from visual to auditory, and even haptic feedback. Modern examples tend to include high budget games that require multiple senses needed to be used to complete the game, such as *The Last of Us*, where the player needs to rely on sound and haptic cues to navigate through the environment, as well as games like *Rhythm Heaven*, where the player must rely on visual and auditory cues to complete the game, with the haptics being used to provide feedback to the player when they make a mistake.

Multimodal videogames can also be designed with specific purposes in mind, being used for educational to health related purposes, for example. These fields benefit from the use of multimodal videogames due to their ability to analyze multiple senses at once, providing a more complete view of the user's current state. This study will focus on the memory aspect of multimodal videogames, specifically on how different external stimuli (video, audio, and haptic) can impact the cognitive load and the performance of the users while they are playing a memory videogame.

1.1 Context and Motivation

Multimodal memory games have been used in various contexts pertaining to the study of their impact on cognitive abilities, such as memory, attention, and problem-solving skills, with games such as *Nintendo Brain Age* and *Lumosity* being proposed by their developers as tools that could help improve someone's cognitive related skills through regular gameplay.

Cognitive load is defined as the amount of mental effort being used in the working memory [19]. This field has been widely studied, but never to the extent of understanding how different external stimuli can impact the cognitive load of users, especially in the context of multimodal memory games.

In the context of cognitive load and the work done in this field, it is possible to find multiple studies regarding the impact videogames have on cognitive load. There are different approaches to measure cognitive load, as seen in published studies, that end up using eye-tracking [10] or pupil dilation [14] to measure the cognitive load of users while they play videogames. However, this field tends to either neglect the impact external stimuli can have on the cognitive load or does not explore the player performance in depth, focusing only on one of these two aspects.

Due to these factors, there is a lack of research that combines both the cognitive load and player performance aspects while also exploring the impact different external stimuli have on these two factors. External stimuli such as audio, visual, and haptic feedback have been used in various videogames to enhance the user experience, but the specific impact of all these stimuli, combined, on these two factors has not been thoroughly investigated.

1.2 Problem

As seen in section 1.1, there is a lack of research regarding the impact external stimuli have on the cognitive load and performance of users while playing multimodal memory games. Little is known about how all these stimuli combined can impact these two factors, and if their impact is significant enough to be taken into consideration when designing a videogame.

External stimuli can have a big impact on the user experience while playing a videogame, be it either by enhancing immersion or by providing important feedback, such as audio cues that guide the player to the objective or haptic feedback to indicate to the player that they have taken damage, for example. Although these factors have been studied individually, as for example, the impact faster or slower music tempos have on player performance [12], how haptic feedback can enhance learning a new instrument [21], or even the difference that watching a show with or without subtitles has on the overall experience [26].

The cases seen previously all involved one unique stimulus, but very little is known about how these stimuli interact with each other when combined, with some researches focusing on specific combinations of two stimuli, such as audio-haptic [5] and visual-haptic [20]. The study of the combined impact of all three stimuli (audio, visual and haptic) on the cognitive load and player performance is an untapped area of research, with still possibilities to explore and expand the knowledge in this field. This could be important not only for the design of multimodal games that use three stimuli, but also for the study of how these stimuli interact with each other, and if they have, in fact, a significant impact while playing a multimodal memory game.

Multimodal memory games can be used in various contexts, such as education or rehabilitation, and understanding how these stimuli can help or hinder the user experience can be important for the design of these games, as understanding how every component works can lead to better experiences for the users, therefore improving their cognitive load and performance in these settings.

As such, there is a clear gap in the literature regarding the combined impact of visual, auditory, and haptic stimuli on cognitive load and player performance in multimodal memory games. This study aims to fill this gap by investigating how all these factors interact with each other, with the goal of providing an answer to this problem.

1.3 Research Questions

Considering the previous statements made in Section 1.2, this study will aim to answer the following research questions:

- **RQ1:** How do different types of visual, auditory and haptic stimuli affect cognitive load and player performance in multimodal memory games?
- **RQ2:** Is the impact on the cognitive load and player performance considerably different with and without combined sensory feedback?
- **RQ3:** How do these three stimuli interact with each other to influence the final results?

Given these research questions, the study will cover variables related to cognitive load and player performance, respectively. These variables will be selected after a careful analysis of the literature, ensuring that each variable is relevant and significant to the study. The variables selected will be further explored in Chapter 3, where it will be explained how they will be measured and analyzed.

It is important to observe each variable independently, in order to understand how each one of them is affected by the different stimulus, and how each stimulus interacts with each other. This will allow for a comprehensive look at the impact each one has on the cognitive load and player performance, therefore allowing to understand their behavior when combined.

1.4 Objectives

To answer the research questions defined in section 1.3, a set of objectives was defined to guide the study and ensure it stays aligned with what is being investigated. For this, each Research Question (RQ) has a set of objectives that will be covered during the study. They are as follows:

- **RQ1**
  - Evaluate the effect of each type of stimulus (visual, auditory, haptic) individually to learn their impact on cognitive load and player performance.
  - Compare the results obtained to understand which one has the most significant impact on cognitive load and player performance.
- **RQ2**

  - Quantify cognitive load and player performance with each stimulus active individually.
  - Quantify cognitive load and player performance with pairs of stimuli active at the same time.
  - Quantify cognitive load and player performance with all three stimuli active at the same time.
  - Compare the results obtained from all scenarios in order to understand the difference between them.

- **RQ3**
  - Analyze the interactions between the three stimuli when combined.
  - Identify which stimuli combinations lead to the most significant changes in cognitive load and player performance.
  - Identify potential pairs of stimuli that may improve or hinder player performance when combined.
  - Compare the results obtained from the combination of all three stimuli with the results obtained from individual stimuli and pairs of stimuli analysis.

Achieving all these objectives will provide valuable insights into this field of study, helping fill the gap previously identified in section 1.2 and providing a solid foundation for future researches.

1.5 Contributions

This dissertation aims to provide an analysis based around the multimodal memory game field, including the cognitive load and player performance aspects. Knowing that, the contributions of this dissertation are as follows:

- A multimodal memory game that integrates three different types of stimuli (visual, auditory, and haptic) across seven different conditions (including individual stimuli, pairs of stimuli, and all three stimuli combined).
- Results show that, when alone, the visual stimulus has the better results in terms of cognitive load and player performance, while the haptic stimulus has the worst results.
- Combination of stimuli does not significantly improve the results obtained from the visual stimulus alone, but it does slightly improve the results obtained in some cases, althought almost unnoticeable.
- The framework developed for this study can be used in future studies, as it supports dinamic changes in the stimuli used, allowing for an easier integration of new levels.

1.6 Dissertation Structure

This document is structured as follows:

- **Chapter 2 - State of the Art:** This chapter provides an overview of the existing literature and research related to the theme of study in this dissertation. It covers relevant topics in these fields and highlights ideas and concepts that may be useful in the development of the study. The literature review can also be found on this chapter, providing a clear and concise summary of how the research of articles and documents in this field was conducted. In this chapter, the research methodology used to search and select all the relevant articles used to support this study is also presented, while naming the databases and repositories used in the research done.

- **Chapter 3 - Approach for Measuring Cognitive Load and Player Performance:** This chapter introduces the methodology and work developed to measure the cognitive load and player performance. It covers the design and implementation of the multimodal memory game used in the study, including how the stimuli were integrated and how data is collected. The conceptual design and architecture of the developed project are also presented, giving insights into the technical side of the work.

- **Chapter 4 - Evaluation and Results:** This chapter presents the experimental setup for evaluation and the results obtained from the evaluation of the multimodal memory game, including the analysis of the data collected regarding cognitive load and player performance. The evaluation covers the order of tasks and the metrics collected for the study, while the results obtained are discussed in relation to the research questions and objectives defined in Chapter 1.

- **Chapter 5 - Conclusion:** This chapter summarizes the main findings of the study and the limitations found during development. The possible future work is also discussed, providing insights into how this research can be extended and improved.

Chapter 2

State of the Art

In order to contextualize the area of study and provide a comprehensive understanding of the topics involved in this dissertation, this chapter provides an overview of the state of the art regarding videogames, cognitive science and external stimuli. Each one of these topics is explored in-depth, covering the most relevant concepts and ideas that were collected during the literature review process. The articles and documents here covered were carefully selected from the literature review methodology presented in both Section 2.1 and Appendix A, allowing for a solid foundation to build the basis of this research.

All the topics covered have a direct relation to the study, providing a solid foundation for the research conducted. The chapter starts by exploring the concept of videogames, with a bigger focus on serious games and the differences observed between unimodal and multimodal games. Following that, the chapter delves into the cognitive science field, exploring the concept of cognitive load and how it can be affected by different factors, in here are some works that explore the impact of cognitive load on various fields, such as education and user age, for example. Finally, the chapter explores the multiple external stimuli that can be used to enhance the user experience, combining visual, auditory, and haptic stimuli with the previous fields studied.

With the knowledge obtained from this chapter, it will be possible to move forward with the research, by identifying possible gaps or future work that can be done in this area, observed through the studies presented and analyzed. Identifying these gaps will allow for an easier definition of what can be done in order to further research these fields, as well as helping define the methodology and design used for the future proposed solution.

2.1 Research Methodology

Before delving into the state of the art, it is important to understand how the articles and documents that will support this study were obtained. A reliable research strategy was put in place, including various different steps to ensure that the articles obtained were relevant and from reputable sources.

To obtain the articles used in this chapter, a set of databases and repositories were used, including Google Scholar, IEEE Xplore, ACM Digital Library and Scopus. These repositories were chosen due to their vast collection of articles, which is necessary for a complete and reliable literature review. Each repository has its advantages and disadvantages, but when put together, they provide a good coverage of the topics related to this study.

Some other conferences that didn't show up in the repositories research were also explored, such as the NIME: New Interfaces for Musical Expression, for example. This extra search outside of the aforementioned repositories was necessary to ensure that, in case some relevant conferences weren't covered by the repositories, they would still be included in the literature review process.

In Appendix A, a more in-depth look into the research methodology used will be presented, including the search queries used, the inclusion and exclusion criteria and the total number of articles obtained from the literature review process.

2.2 Videogames and Their Applications

Videogames have been around since the early days of computing, evolving from simple pixelated games to complex and immersive experiences. Over the years, videogames became the most popular form of entertainment, becoming one of the most profitable industries worldwide, with the most recent balances reporting the industry being worth over 184 billion USD in 2024 [11].

This exponential growth has led to the development beyond entertainment, extending their use to other fields such as education and healthcare, for example. This allowed for the emergence of Serious Games (SG), as well as the exploration of different sensory stimuli on said games, which could enhance the overall user experience.

2.2.1 Serious Games

SGs are designed with a focus beyond entertainment, their main purpose being to educate, train or raise awareness about specific topics. These games can be defined in a variety of ways, with possible definitions going from "*a promising technological approach to cognitive assessment because they make assessment entertaining and easy to perform, and they enable a more cost-effective and frequent cognitive assessment.* [27], ultimately showing that a SG does not need to sacrifice fun in order to be effective. Another approach points out similarities between SGs and methods of cognitive stimulation used in these studies, "*Serious Games share some characteristics with other methods of cognitive stimulation ... these games involve a task which is continuous, systematic, stimulating and providing reinforcement, avoiding the routine and repetition. The task can also stimulate specific cognitive domains, especially attention-concentration, facilitating the stimulation of other domains.*" [18], showing that SGs can be designed in order to target specific cognitive domains, making them a great tool for cognitive training and assessment.

In this field of research, promising results regarding the use of serious games for cognitive load measurement, including player performance, have been observed and documented, allowing for a deeper understanding of how these two concepts interact with each other. It has been discussed that the impact of these games on their users can be measured through different methods, with some more effective than others. One of these methods is pupillometry, which measures the pupil dilation as an indicator of cognitive load. This method has been used with positive results, as pupil diameter shows a significant difference between performance levels, with lower performing players showing longer latency times before reaching peak pupil dilation, suggesting a slower cognitive processing speed, due to the increased cognitive load experienced, as seen by Calà et al. [4].

2.2.2 Unimodal and Multimodal Cues

Games, including SGs, can be designed to use different sensory stimuli to enhance the user experience. These stimuli can be classified into unimodal and multimodal cues. Unimodal cues, as the name suggests, involve the use of a single sensory modality, such as auditory, visual or haptic stimuli [3]. On the other hand, multimodal cues are the combination of two or more of these sensory modalities, aiming to create a more immersive and engaging experience for the player [3], and support can derive from a number of different human factors, such as gestures, touch, vision and haptics [28].

On the multimodal side, it has been observed that these cues have better performance when put directly against unimodal cues. Studies show that multimodal cues can lead to better overall performances in a variety of tasks, such as the study of the impact vibrotactile and auditory feedback have on performance in a search game [2], where it was proven that the combination of both variants did in fact lead to better performance from the participants. This could help support the idea that playing, for example, a tabletop game with multimodal cues could lower the cognitive load of the players and be perceived as more helpful than other approaches, as seen by Zimmerer et al. [28], further supporting the idea that multimodal cues are indeed the best.

But there is more evidence supporting this concept, as presented by Chin et al [6], who investigated the impact of multimodal cues on learning efficiency and performance in a musical game. The system created for this study combined visual, auditory, and haptic stimuli to provide feedback to the players. The results showed a great improvement in learning efficiency with this combination, with an increase of 31.1% being reported by the authors. The authors suggest that the multimodal approach helps reduce uncertainty, which can be seen as a reduction in cognitive load, due to the different sensory inputs reinforcing each other.

Although the majority of studies point towards multimodal cues being the better approach, unimodal cues can still have their place in certain scenarios. When directly comparing two distinct unimodal cues on a simple task, such as haptic and auditory, it has been proven that either one of these cues can be enough to provide the necessary feedback for the user to complete their task [3].

With these concepts in mind, it is clear that both unimodal and multimodal cues have their place in game design, with multimodal cues being the better option in most scenarios. However, unimodal cues can still be effective and less cumbersome in simpler tasks, where the cognitive load is already low enough for the user to perform well without additional stimuli. Knowing when to use each approach is crucial for the future work that will be done in this research, as it will help define in which scenario each cue works better.

2.3 Cognitive Load and External Stimuli

Cognitive load is a concept from cognitive science, which is the field that studies the mind and all its processes, including how people think and learn, their psychology, philosophy and neuroscience, besides many others [16]. Inside this broad field, one of the most relevant concepts for this study is cognitive load, part of the Cognitive Load Theory (CLT).

Once there is a consensus of how CLT works, it is important to explore the impact cognitive load can have on various fields. This section will explore some studies that focus on this impact, highlighting the importance of understanding cognitive load and its implications.

2.3.1 Cognitive Load Theory

According to studies CLT, cognitive load is the amount of resources used by working memory for a specific task [19]. This theory is based on the fundamental structure of working memory and long-term memory. For example, long-term memory has a limit to the amount of information it can store, but it is also responsible for the majority of information retention and cognition, while working memory manages the conscious and structured processing of information [19].

With this in mind, CLT can be divided into three different types of cognitive load: intrinsic, extraneous and germane cognitive load [9, 15, 25]. Intrinsic load is the inherent difficulty of the task at hand, taxing working memory based on the complexity of the information being processed, but high intrinsic load can overwhelm the working memory, which could lead to errors and decreased performance [15]. Extraneous load, on the other hand, comes from how the information is presented to the user (insufficient presentation, bad or disorganized information, etc), which is often caused by cognitive fatigue and lower performance provoked by cluttered environments [15]. Finally, germane load focuses on learning and schema development, often engaging both working memory and visual attention [15].

Recognizing how all these different types of cognitive load interact and affect each other is crucial when designing tasks and environments that aim to optimize cognitive load for better performance and learning [9]. For example, mitigating the complexity of tasks related to intrinsic load and refining the ones related to extraneous load can lead to better germane load, which suggests that a serious game can be designed with this in mind, reducing unnecessary cognitive load and enhancing germane load for better learning outcomes, as studied by Choi et al. [9].

These interactions can also be observed in other contexts, such as videogames. Games like *Ori and the Blind Forest* offer the player a vast world for them to explore and interact with, although with a fair share of frustrating moments that could hinder the player's experience and performance. These moments could be seen as an obstacle to the player, possibly making them want to quit the game, but in reality, these encourage player adaptation and engagement in the long run, due to the right balance between intrinsic and extraneous cognitive load, which in turn supports germane load [25]. Although this effect can be seen in this game in particular, it can be extrapolated to other games as well, as long as the right balance between these different types of cognitive load is achieved.

2.3.2 Cognitive Load on Various Fields

With the understanding of CLT and its different types of cognitive load, it is important to explore how these concepts can be applied to various fields, such as education and user age. Various studies have been done with the purpose of understanding how cognitive load can impact each individual person differently, depending on their age or current health status.

There are a lot of different ways to study cognitive load. Some of these have been mentioned previously in this dissertation, such as pupillometry [4], but there are many others, such as heart rate variations [1], changes in time, such as day and night cycles [23], and different groups of people of different ages [18] and even eye-tracking activity [15].

Let's take the example of age related cognitive load. We can take in consideration that as people age, their cognitive abilities tend to decline, which can have impacts on their day to day activities, like decision making and memory retention. This however, can be mitigated through regular cognitive activity with moderate intensity, as pointed out by Neto et al. [18], through the use of simple interfaces and tasks that aim to reduce unnecessary cognitive load, leaving more resources available for the more important and focal tasks. Knowing this, these changes of simplifying interfaces and tasks could also be applied to videogames, especially those designed for brain training, such as SG, as these have been shown to have positive effects on older adults, being able to maintain and improve their attention and memory skills, avoiding further cognitive decline [22].

But this is not the only field where cognitive load can have a significant impact, it is even possible to observe changes in cognitive load based on the time of day, or to be more specific, day and night cycles. This impact can either be observed in real life scenarios or even in videogames, with both having the same results, independent of the environment, a person will have a higher cognitive load during night settings, leading to worse performance across all tasks they try to perform [23]. This could be justified by the natural behavior of the human body, as it is designed to be more active during the day rather than at night, however this assumption would need to be further explored. This can be observed on 2.1, where the Task Completion Time (TCT) is higher during night settings, regardless of whether the task was performed during while it was day or night in real life, night settings always showed similar or higher TCT values when compared to day settings.

More importantly, cognitive load can also be explored through behavioral and psychological methods, with these two methods being capable of predicting cognitive load reliably in various scenarios, independent of each other. One of such scenarios could be in emergency situations, where cognitive load can have a significant impact on the final outcome if the situation is not handled properly. To measure these effects, and make the results more reliable, non-intrusive data measures need to be used, as intrusive methods, such as Electroencephalography (EEG), could interfere with the results, due to the cumbersome nature of these devices. Methods such as pupillometry, eye-tracking and heart rate are all effective ways to measure cognitive load without breaking immersion and presence. With these methods, we can return to the emergency situation

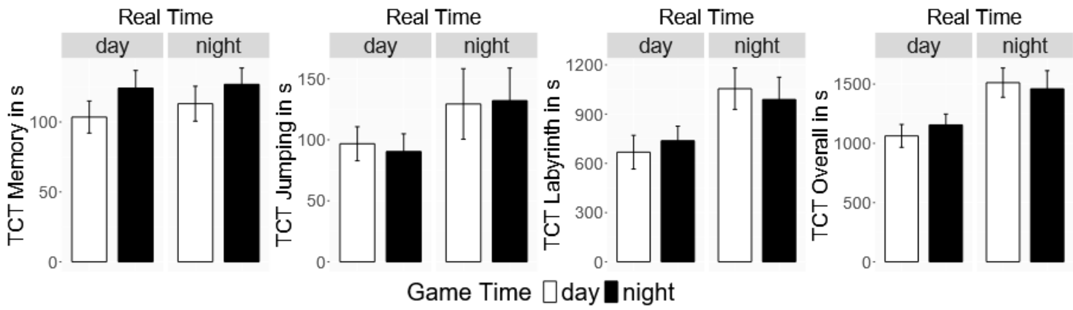


Figure 2.1: TCT during day and night settings.
[23]

example, and, as shown by Appel et al. [1], who used these previously mentioned methods in their study, we can in fact see that cognitive load does in fact increase with task difficulty, such as in emergency situations. This study not only was able to prove this concept through the methods used, but also showed great accuracy in its results, which could justify the use of these non-intrusive methods in future studies regarding cognitive load.

These studies highlight the importance of understanding cognitive load and its implications across various fields. By recognizing how cognitive load can impact individuals and situations differently, we can design better tasks and environments that optimize cognitive load for improved performances and learning outcomes. Combining the knowledge obtained from these studies with the concepts learned about CLT can enhance our understanding of these designs and further prove the development of this research.

2.3.3 External Stimuli

External stimuli are the sensory inputs that users receive from the environment. The stimuli can be either visual, auditory, haptic or a combination of two or more of these. These stimuli play a crucial role in shaping user experiences, especially in interactive environments such as videogames, as they help create a sense of presence and immersion for the user. Each stimulus can have a different impact on the user, depending on how they are designed and implemented. Understanding how these stimuli work individually and in combination is essential to create the best possible experience for the user, and on understanding how they can impact cognitive load.

Visual stimuli is often the primary sensory input in videogames, providing players with essential information about the game world and its objectives, either by introducing visual elements that catch the user's attention or by limiting excessive visual cues [10, 15]. Auditory stimuli, while not as dominant as visual stimuli, still play a significant role in shaping user experiences. Sound effects, music, and dialogue can all contribute to the overall environment, enhancing immersion and providing important cues to the player [12, 24]. Haptic stimuli provides tactile feedback to users, being often used through vibration or force feedback, but can also include more advanced technologies, such as wearable devices that provide a more realistic experience.

Knowing how these different stimuli work together is crucial for creating cohesive experiences, without the risk of overwhelming the user with too much information. Combining visual, auditory, and haptic stimuli can create an even more immersive experience, as each stimulus could reinforce each other, possibly reducing uncertainty and cognitive load.

In the next section, these concepts will be further explored, referring to various studies for support of the ideas presented here. While looking into these advances in the field, the limits and gaps that still exist in this field of study will also be identified, allowing for a better understanding of what can be done to further research these topics.

2.4 Works Related to External Stimuli

External stimuli play a crucial role in shaping user experiences, especially in interactive environments such as videogames. These stimuli help create immersive worlds that engage multiple senses, enhancing the overall experience. This section explores such external stimuli, focusing on visual, auditory, and haptic stimuli, as well as their combined use.

These stimuli can be directly connected to the previous topics covered in this chapter, as they can also have an impact on the cognitive load of users, be it either positive or negative. Understanding how all these interactions work together is an important step towards creating a better designed environment for users.

2.4.1 Visual Stimuli

As presented in the previous section 2.3.3, visual stimuli is often the primary sensory input used in videogames, providing players with the utmost essential information about the game world. The impact of this stimulus on cognitive load has been explored in-depth before, giving the possibility to understand how this stimulus actually affects users.

One could say that the tracking of this component of cognitive load is still in its infancy, but it has been shown that a simple eye-tracking system is enough to provide valuable insights into this stimulus. With the use of eye-tracking, it's possible to observe that visual stimuli used as visual guidance can support performance but also introduce additional attentional demands, which could lead to an increase in cognitive load [10], and as indicated by Figure 2.2, where the gaze as concentrated in one place only with the assistance of visual stimuli, contrary to the heatmap for no visual stimuli, where the gaze is more dispersed. With this in mind, it's important to consider that while designing and introducing visual components into an environment, the developer should be careful not to overload the user with excessive visual information, as this could lead to an increase in cognitive load, which in turn would make the user perform worse in their tasks, as seen previously in 2.3.2. While this study by Fritz et al. showed promising results, further development can still be done in order to obtain more accurate results, such as individual eye calibration for each user, due to each person showing different eye behavior and pupil dynamics.

(a) Heat map without assistance (b) Heat map with virtual flight director (c) Heat map with ELA on the tablet

Figure 2.2: Gaze heat map with and without visual stimuli assistance. [10]

These findings could be extrapolated to other interactive environments, such as videogames, including Augmented Reality (AR) and Virtual Reality (VR). One such case is the use of this technology in SGs, with a pertinent example being related to the use of these videogames for training of highly intensive tasks, such as emergency landings done by pilots for example [10]. For these SGs, good visual guidance is crucial, as a small difference in details could hinder the performance of the user, affecting their learning outcomes and potentially destabilizing the cognitive load during these tasks.

While this field is well documented, there is still room for improvement. As mentioned previously, taking into consideration individual eye calibration is an important step towards obtaining more accurate results. The display itself could also have an impact on the results, as different displays have different characteristics, as for example brightness and contrast that could impact how the user perceives the visual stimuli [20]. All these factors need to be carefully analyzed and studied before implementing visual stimuli in the videogame environment, as anything that hinders the visual effect could affect the user and its cognitive load negatively.

2.4.2 Auditory Stimuli

Auditory stimuli can be applied in various ways in videogames, from background music to sound effects. The use of background music and tempo variations can influence player performance, as variations in music tempo can either assist or disrupt the performance of players, depending if the tempo matches the pace of the game or not. This can be observed in the videogame *Tetris*, where faster tempos increased the cognitive load related to action synchronization, while an asynchronous tempo had the opposite effect [12]. This fact could also be used to personalize game difficulty, by adjusting the tempo to match with the player skill level, with faster tempos being used for more skilled players, while slower tempos could be used for players who are just starting out.

But having auditory stimuli present in videogames does not automatically lead to a better experience for the user. Sometimes, this stimulus could have no significant impact on the user experience, meaning that adding them would only lead to an unnecessary increase in cognitive load [24]. This highlights the importance of properly designing and integrating auditory stimuli into the game mechanics, they should only be added when they provide meaningful information or to enhance immersion, otherwise they wouldn't be necessary.

This field of study, contrary to visual stimuli, is still rapidly evolving, with a lot of untapped research potential still unexplored. Knowing which auditory stimuli work best in different scenarios is crucial for the future work that will be done in this research, as well as understanding how these can impact the player experience and cognitive load.

2.4.3 Haptic Stimuli

Haptic stimuli, as seen previously, provides tactile feedback to users, often through vibration or force feedback. This stimuli is used to enhance the sense of presence and immersion in videogames, providing a more realistic experience to the user. The impact of this stimuli on cognitive load and player performance has had mixed results, with both positive and negatives effects being reported.

This stimulus has been shown to be an effective alternative to other stimuli, such as visual and auditory, in certain scenarios. If properly designed, haptic feedback could also complement these other stimuli, with the possibility of enhancing the overall experience [8]. On the other hand, haptic stimuli can also lead to an increase in cognitive load, worsening the performance of the user [13]. This was observed in a VR environment, with authors believing this worsened performance, being related directly to the unfamiliarity of the users with haptic feedback, advising it should be used alongside other stimulus, and the fact that the hand tracking used was not optimal, leading to latency issues and lower tracking accuracy. This could potentially be mitigated with the use of a controller, or even the research if this behavior persists in more familiar environments, like a SG without the use of VR.

However, haptic stimulus can also have a positive impact on user experiences. Taking the example of musical interactions, it has been documented that introducing haptic feedback into musical tasks, such as playing or learning an instrument can enhance the overall experience, providing a better understanding and reducing uncertainty for the user [21]. It is suggested that applying this same logic to SG can also enhance learning outcomes, as well as the combination of haptic with other stimuli for learning purposes.

With these results, it is clear that haptic stimulus can have mixed results on user experiences, depending on how it is used and designed. Multiple studies show that this stimulus should be accompanied by other stimuli, such as visual and auditory, in order to provide a more cohesive experience and avoid overwhelming the user. Another important factor to consider is the familiarity to this stimulus. If a user does not have experience with it, introducing haptic in for example a VR environment could lead to an increase in cognitive load, while if they are in a more welcoming environment, such as a SG designed for learning, haptic could be used more effectively.

2.4.4 Combined Stimuli

While all the previous sections explored each sensory stimulus individually, it is important to understand how these stimuli can interact with each other when combined. Combining different stimulus can lead to a more controlled and defined experience, as each stimulus can reinforce one another, reducing uncertainty and helping manage cognitive load.

Combining visual and haptic stimuli has been shown to have more benefits in capturing the user's attention, more than just visual stimuli alone, with a special focus on cognitively demanding tasks [20], confirming the idea proposed previously in 2.4.3. Visual stimuli alone can increase cognitive load, especially when the user is already visually overloaded with information. The addition of haptic stimuli could help direct attention more effectively, in a way that reinforces the visual cues provided. This, however, would need to be properly integrated and taken into consideration, as mentioned in 2.4.1, every individual has different visual behavior, and what could work for one person might not work for another.

On a similar note, auditory and haptic stimuli combined have shown mixed results, such as the addition of gestures in a VR environment being excessively demanding for the cognitive load when combined with audio interpretation [5]. With the removal of gestures, the cognitive load was reduced significantly, confirming again what was seen previously in 2.4.3.

Combining stimuli can be used in various scenarios, such as music learning. Sometimes, while playing an instrument, audio might not be enough to provide the necessary feedback to improve performance. In these scenarios, adding another stimulus, either a video on a screen showing the note played or haptic feedback notifying the user of a wrong note, can help reinforce learning and reduce uncertainty. This was observed by Chin et al. [7], who proved that the addition of haptic does in fact increase performance and learning outcomes.

These results not only show that combining different sensory stimuli can enhance user experience and help with cognitive load, but they also need to be carefully pondered and implemented, as a poor design could lead to the opposite effect of the one intended, increasing cognitive load and hindering performance.

2.5 Summary

The knowledge obtained from this chapter will be used to support the development of a SG that properly integrates visual, auditory, and haptic stimuli aiming to learn how this combination can impact cognitive load and player performance. During development, the research will try to aim for some of the errors and gaps observed, and knowledge obtained from the studies analyzed will also be taken into consideration, such as using controller haptics instead of hand tracking, offering different difficulty levels to mitigate intrinsic cognitive load and offering different auditory tempos to match the player's chosen difficulty.

Going over the videogame section, it is clear that interactive environments evolved significantly over the years. What started as a way of entertainment expanded to other fields, such as education and research, leading to the emergence of SG. These games have been shown to be effective tools to measure cognitive load and player performance, and with the correct use of unimodal and multimodal cues, these videogames can provide a more in-depth study of these concepts.

Cognitive science provided the necessary framework to understand cognitive load and how it works. With this understanding, it was possible to explore the impacts cognitive load have on various fields of study, and the impact external stimuli can have on this concept. Keeping in mind the different types of cognitive load, we can plan a better design for our study, as for example, we can aim to mitigate intrinsic load and focus more on extraneous and germane load, which will reduce unnecessary cognitive load and enhance learning outcomes.

For the external stimuli section, it is clear that all these stimuli can have a significant impact on the user and their cognitive load, with more positive results than negative. Knowing this, developing a SG that properly uses all these stimuli combined is a promising approach to futher research in this field of study. All studies analyzed showed a combination of two different stimuli or only one stimulus being used, but none explored the combination of all three, with the outcome of this combination being unknown.

The next chapter will explore the methodology and design used to develop this SG more in-depth, discussing the approach and conceptual overview of the proposed solution, including how it will be experimented with and evaluated. A proposed timeframe will also be presented, showing the different steps that will be taken for this research.

Chapter 3

Approach for Measuring Cognitive Load and Player Performance

In this dissertation, we propose an analysis of the impact that different types of stimuli have on a given user when interacting with a memory-related video game. For this purpose, a memory digital game was developed with the goal of observing and collecting data around this subject, with the end goal of understanding the differences between the stimuli used and the cognitive load they generate on the player, as well as the impact on their performance.

In this chapter, we will describe the project developed and the methodology used to collect the data needed, including the equipment used to integrate the different stimuli into the game and how this implementation was conducted.

3.1 Requirements

To measure cognitive load and player performance, we needed to design a game that could integrate all stimuli in a way that would not overwhelm the user or feel too intrusive, while still being able to measure the impact of each stimulus on its own, as well as the impact of mixing and matching them together. This meant that the game had to be simple enough in order to not generate outside noise that could interfere with the measurements, while also being engaging to keep the players' attention and motivated to play.

To do this, a memory game was developed based on a simple and minimalistic design, allowing the user to focus on the stimuli and level presented without being distracted by other elements, such as the background, for example. Outside noise is a common problem in studies based around cognitive load, as any element introduced in the game can have an impact on how the player performs, be it either positive or negative. Such impacts can then give a false representation of how the stimuli actually impacted the users, as they could have been distracted by unimportant elements in the game, or used them as a support tool to perform better.

To avoid this, the game was designed with the following requirements in mind:

- The game should have two different types of tasks, one based on the "Simon Says" game, where the player has to repeat a pattern shown through the stimuli, and another based on the "Memory Game", where the player has to match pairs of cards from a grid.
- Must be played on a touchscreen, which in turn allows the removal of peripheral devices such as a gamepad controller or a mouse, which could have generated additional noise in the results obtained, in case the player was not familiar with these devices.
- To minimize bias in the results, each participant should play the levels of a task in a randomly assigned order, defined before the experiment starts. If every participant plays the same order of levels, the order itself may influence the results. For example, participants may perform worse on the first level as they are still learning how to play the game, in turn, leading to better results on later levels, as they become familiar with the game. Alternatively, participants may also become tired or lose focus over time, resulting in a worse performance on the later levels. In either case, these differences could be caused by the order the levels are played, instead of the levels themselves. Randomizing the order helps reduce these possible biases and ensures more reliable results.
- Also, to avoid bias, as the game is divided into two different types of games, the order these are played first should alternate between each other. For example, if an user starts with the "Simon Says" levels, the next one should start with the "Memory Game" levels, and so on.

Besides these design requirements, there is also the need to define the different levels of the project and how they will be structured, which data will be collected for future analysis, and which external equipment will be used to collect such data and how it will be integrated into the game. These topics will be covered in the next sections of this chapter.

3.2 Conceptual Model

The conceptual model of the developed project is based on the idea of a self-contained project that can control itself independently, be it either from internal components like the game logic and structure, or from external factors, such as the equipment used to introduce different stimuli into the gameplay. We can divide these components into two different blocks: the components needed for the game's functionality, and the components needed for the data collection.

Figure 3.1 illustrates the conceptual model of the project, where it is possible to see the different interactions and components involved in the work. The components represented by a continuous line are needed for the game's functionality, while the components represented by a dashed line are needed for the data collection. The components are as follows:

- **Session Manager** - the core of the project, managing the flow of the player activity, including the order of the levels played, and the integration of external components for data collection.

- **Simon Says Manager** and **Memory Game Manager** - share the same behavior, as they are responsible for managing the game logic and structure of their respective games, while also controlling which stimuli is currently active in a given level.
- **Haptic Manager** - pertains to the external component used for the haptic stimuli, in this case, an Arduino board directly connected to a vibration motor.
- **Eye Tracking Manager** - responsible for the integration and control of the eye tracking device used, in this case a *Gazepoint GP3 HD*.
- **Questionnaire Manager** - responsible for the integration of the questionnaires used to evaluate the cognitive load of the player, in this case the NASA Task Load Index (NASA-TLX) questionnaire, which in this project is implemented through the use of *Google Forms*.
- **Data Manager** - responsible for aggregating and structuring all the data collected from the different components.

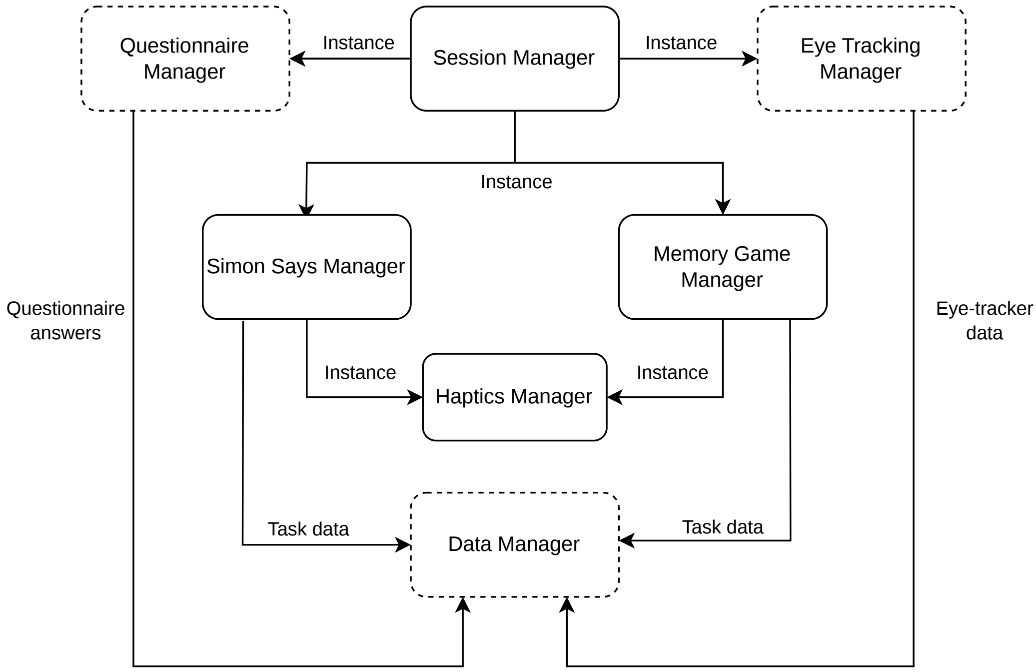


Figure 3.1: Conceptual Model Diagram

On the game functionality side, we have the **Session Manager**, the **Simon Says Manager** and the **Memory Game Manager**, and the **Haptic Manager**. These components are responsible for the game logic and structure, ensuring that the game is playable and that the different stimuli are properly integrated into the gameplay. It is also important to highlight that. Originally, the **Simon Says Manager** and the **Memory Game Manager** were combined into a single component, as the original idea was to have a single self-contained game that would integrate both types of tasks as the participant progressed through the game. However, after further analysis, it was decided to separate these two components, not only because their logic is different, but also because it allows for a more flexible approach to the design of the game.

Meanwhile, related to the data collection, we have the **Eye Tracking Manager**, the **Questionnaire Manager** and the **Data Manager**. These components are responsible for collecting the data not integrated into the game itself, but instead collected through external components, such as the eye tracking device and the NASA-TLX questionnaire. Everything collected is treated by the **Data Manager**. The **Data Manager** being treated as an external component is a design choice, as it is not integrated into the game itself, it is only used after the game is completed, saving the data collected without interrupting the participant's experience.

These components will be further explained in the next sections of this chapter, allowing for a better understanding on how they work and how they behave in this setting.

3.3 Architecture

The architecture of the project is based on the previous conceptual model, where each component is now divided into different modules and connected together, in order to satisfy the requirements needed to conduct this research. The final application was built using the *Unity* game engine, on version *6000.3.9f1*, a *Gazepoint GP3 HD* eye tracking device, and an *Arduino Leonardo* board for the haptics used in the game. The final build is then tested and played on a touchscreen monitor, making it more accessible to a wider range of participants, as it avoids external peripherals like gamepads or mice, which could be confusing for those unfamiliar with them, therefore impacting the results.

Figure 3.2 illustrates how the various components are connected and share data with each other. The figure shows that the user interacts with the game through a touchscreen monitor, has their gaze and pupil dilation collected through the *Gazepoint GP3 HD* eye tracker mentioned previously, and is also asked to fill out a questionnaire, contained inside the *Computer* container, once the game is completed.

The touchscreen monitor is connected to the *Computer* component, same as the *Gazepoint GP3 HD* eye tracker, but one is linked to the *Application* container, while the other is linked to the *Gazepoint Software* container, respectively. In the *Application* container, we have every script responsible for the game logic and structure, as well as the haptic stimuli integration, that can either be triggered or not by the game logic, depending if the level being played is using haptics or not. Once the game is completed, the metrics collected from the game is sent to a *.json* file, containing mainly data related to the performance of the player. On the other hand, the *NASA-TLX* questionnaire also becomes available, allowing the participant to fill it out on the same computer the application was run. Once the participant fills out the questionnaire, the answers are saved in a *.csv* file, which is then used for future analysis.

The *Gazepoint Software* container is responsible for controlling and collecting the data captured by the eye tracker. This software is provided by the manufacturer of the device, and is run

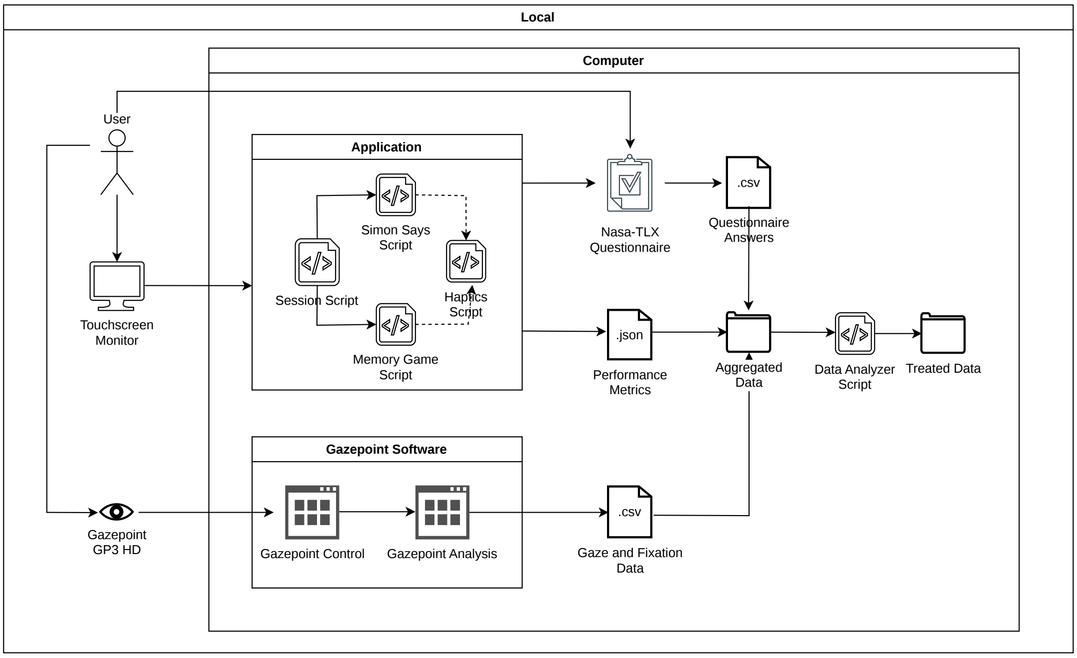


Figure 3.2: Architecture Diagram

on the same computer as the application, in parallel with the application itself. The data collected is then saved in a *.csv* file, to be later used to evaluate the cognitive load of the participant.

With the metrics collected from the game, the eye tracker and the questionnaire, it is possible to filter and structure the data by participant, through the use of a unique identifier generated at the start of the experiment. The data is then cleaned and synchronized, through the use of a *Data Analyzer* script, which is also responsible for analyzing and evaluating the data collected, generating the final results of the experiment.

In sum, this architecture contains the containers and components needed to run the experiment, with each container being self-centered and independent from each other. The evaluation of the data collected can only be done after both the *Application* and the *Gazepoint Software* containers are completed. This way, we ensure that the evaluation of the data is done only after everything is collected, and that the data is not evaluated with missing components.

3.4 Proof of Concept

The initial development of the project was focused on creating a 3D environment where the player would be able to explore and interact with different objects, whose interactions would trigger different puzzles to be solved. These puzzles would be based on the "*Simon Says*" and "*Memory Game*" mentioned before, and the different stimuli would be integrated into the puzzles. The participant would be able to interact with the environment through a first person perspective, using a gamepad controller to move around and the triggers to interact, with these triggers also being used to select different options in the puzzles themselves, due to the resistance they provide, which could be used as a way to implement haptics into gameplay. However, a few issues were raised with this approach:

- How can we guarantee that the player will be able to understand the mechanics of the game? With a more complex game like this first idea, there is a risk that the participants will not be able to understand how to interact with the game, which in turn could impact how they see and complete the puzzles, therefore impacting the results obtained.
- What if a participant is not familiar with the controller used? People have different levels of familiarity, while someone might be used to playing with a controller, others might not know how to use it. It is necessary to find an approach accessible to everyone, to avoid possible bias in the results obtained.
- The current approach is more focused on exploration and interaction with the environment, which could impact the results obtained from the puzzles, as the participants could be distracted by the environment itself, or use it to their advantage in the puzzles. We need to find a way to avoid this, and make sure the puzzles and their stimuli are the main focus of the project, with no outside noise that could interfere with the experience.

To solve these issues, the project was reworked and redesigned, now with the goal of creating a simpler and minimalistic approach to the case at study, as described in section 3.1. First, the user input was changed to a touchscreen, which, when compared to a gamepad controller, is more intuitive and accessible to a wider audience, as it does not require a high level of familiarity with the device as for example a mouse or the aforementioned controller. A touchscreen is also more commonly used, and therefore well versed by a wider range of people, as it has been used in mobile devices like smartphones and tablets for a long time. These conditions make it a more suitable choice for the project, as it reduces the risk of bias in the results to a minimum, while also making the project interactive and engaging for the participants.

It is also necessary to highlight how we implemented the different cognitive load types seen in chapter 2 of this dissertation. Intrinsic load is controlled by keeping the complexity of each task constant across conditions, while extraneous load is minimized through the use of minimalistic interfaces and edited audio clips, as we will see further in this chapter. Germane load is the different stimuli used in the project, as each stimulus condition aims to support different memory enconding strategies, as we will see by the two different types of games used in the project.

In this study, intrinsic load was controlled by keeping task complexity constant across conditions (same sequence length in Simon Says, same grid size in Memory Game). Extraneous load was minimised through minimalistic interface design and carefully edited audio clips. Germane load was the primary variable of interest, as each stimulus condition aimed to support memory encoding differently

While the input method was solved, the redesign of the project itself was not. A complete redesign was now needed to fit with the touchscreen device, as the current approach is still too complex for the study. To solve this, the project was redesigned to be a 2D game, divided into two different types of games, each one with its own levels, but keeping the same mechanics and stimuli, as we will see in section 4.1.2. This approach not only allows for the simple and minimalistic design we needed, but it also makes the participants focus on only the puzzles and their stimuli, as all outside noise was removed, such as the exploration and interaction with the environment. This way, we can guarantee that the results obtained are more accurate than the first conceived approach.

With these changes, the different levels were developed and implemented. For each type of game, seven different levels were created, with each level designed with one or more stimuli, with the goal of observing how the user reacts to each one, and if they perform better or worse with a certain type of stimulus present. Each level has the stimuli it uses indicated before the player starts it, as shown in the Figure 3.3. The order of the levels is also randomized for each user, to avoid bias in the results, as mentioned in section 3.1.

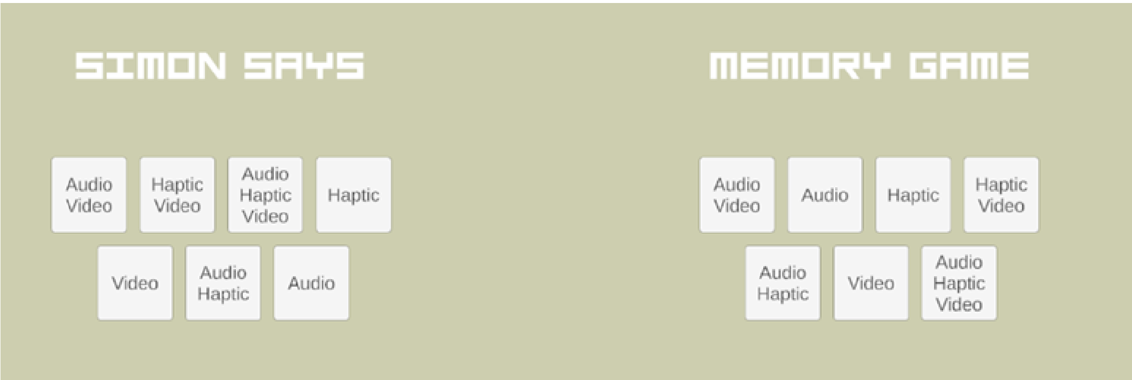


Figure 3.3: Level Order and Which Stimuli are Being Used

As seen on Figure 3.3, the set of levels on the left side represents the levels of the "*Simon Says*" game, while the ones on the right represent the levels of the "*Memory Game*". These two games follow their own logic and structure, but end up implementing the same stimuli. The reason for having two different types of games is to be able to observe the results obtained in two different contexts, as a participant might perform better with a certain type of game than with the other, be it either from being more familiar with one of the games, or because they find one of the games more engaging than the other.

For the "*Simon Says*" game, when the user clicks to start a level, a pattern is shown through the use of different stimuli, and the user has to repeat it by selecting the correct options shown on the screen. These patterns are randomly generated for each user, and if the user selects the wrong option, a brand new pattern is shown, with the same length as the previous one, but with different options. Although a mistake restarts the level, this level only restarts when the user selects a button dedicated to that, which only appears after the user selects a wrong option, as seen in Figure 3.4.

On top of these mechanics, the different stimuli are implemented in the following way:

- **Visual Stimulus** - The visual stimulus is implemented through the use of small flashes inside each square when the sequence is playing. These flashes are used to indicate the order of

Figure 3.4: A level of the "Simon Says" Game, after the user selects the wrong option

the sequence, therefore being used to alert the user to them, supporting the memorization of the sequence this way.

- **Auditory Stimulus** - The auditory stimulus is implemented through the use of different sounds played when the sequence is playing. These sounds are directly related to the objects inside each square, for example, a bell sound is played for the square with a bell icon, a chicken cluck sound is played for the square with a chicken icon, and so on. These sounds are distinct and easily recognizable, so they can be easily associated with the objects in the squares. The auditory stimulus is also used for feedback, as a correct selection will trigger the sound of the object selected, while a wrong selection will trigger a distinct sound, which is the same for all the squares.

- **Haptic Stimulus** - The haptic stimulus is implemented through the use of vibrations, which are triggered when the sequence is playing. These vibrations have a similar behavior to the audios seen on the auditory stimuli, as they are also directly related to the objects in the squares, with different patterns and intensities for each one. Besides these vibrations, a distinct vibration is also triggered when the user selects a wrong option, as we saw on the auditory stimuli.

On the other hand, the "*Memory Game*" has a different structure and mechanics, as the user is presented with a grid of eight cards, with four different pairs of cards, and the user has to find all the pairs by selecting two cards at a time. If the user selects two cards that are not a pair, they are flipped back, and the user has to try again. If the user selects a pair, they are marked as paired, and the cards become unselectable from that point forward. The game is completed when all pairs are found. The different stimuli are implemented in a similar way to the "*Simon Says*" game, with the only difference being in the visual stimuli, which in this case is implemented by showing an image on the card when it is selected. If there is no image when a card is selected, that means the level does not contain the visual stimuli. The auditory and haptic stimuli have the same behavior as in the "*Simon Says*" game, with no differences between the two games.

Figure 3.5 it is possible to see a representation of a level of the "*Memory Game*", where it is possible to see the grid of cards and how cards are marked as paired by a different shade of color.

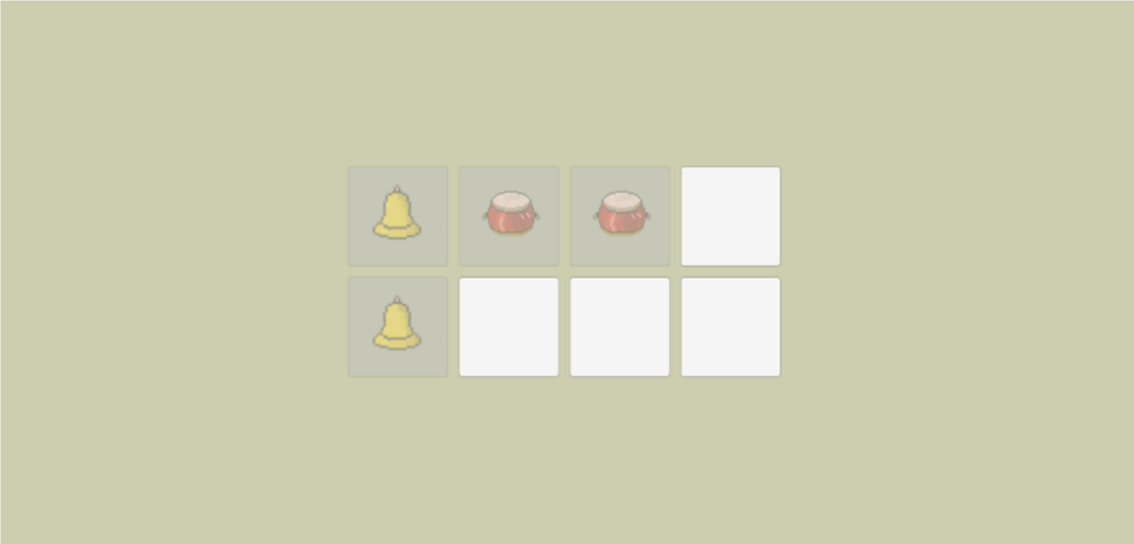


Figure 3.5: A level of the "Memory Game", after the user selects a pair

For both games, the user has a limited amount of time to complete each level, in this case, 60 seconds. If the user is not able to complete the level within this time frame, the level closes automatically, and the user must advance to the next level. This time limit is implemented in a way to give the user a sense of urgency while playing, which can impact the cognitive load of said user. With this time restriction, we can also control the amount of time the user spends on each level, which can be useful to avoid having a player stuck on a level for a long time, or to avoid having a player giving up after a certain amount of time, as they know that the level will end soon, which could push them to keep trying to solve it until the time expires.

Figure 3.4 and Figure 3.5 show some of the objects that can be found inside the levels, which are used to create the different patterns and pairs. These objects were chosen for their distinctiveness and recognizability, as they are all common objects and symbols that are easily associted with their respective sounds and vibrations. The objects used are as follows:

- **Animals** - Chicken, Cow, Horse, Pig
- **Musical Instruments** - Bell, Drum, Guitar, Flute
- **Weather** - Sunny, Rainy, Thunder, Wind
- **Colors** - Purple, Blue, Green, Orange

As seen previously, there are seven different levels for each minigame type, with each level having one or more stimuli. The table 3.1 shows the name of each level and which stimuli are used in each one, with the order the levels show up being unrelated to the name attributed to them. The name of each level is then used to identify the level in the data collection, as we will see in Section 3.6.

Table 3.1: Level Name and Stimuli Used

Level Name Simon	Stimuli Used	Level Name Memory	Stimuli Used
LVL1S	AUDIO	LVL1M	AUDIO
LVL2S	AUDIO + HAPTIC	LVL2M	VIDEO + HAPTIC
LVL3S	VIDEO	LVL3M	VIDEO
LVL4S	HAPTIC	LVL4M	AUDIO + VIDEO + HAPTIC
LVL5S	AUDIO + VIDEO + HAPTIC	LVL5M	HAPTIC
LVL6S	VIDEO + HAPTIC	LVL6M	AUDIO + HAPTIC
LVL7S	AUDIO + VIDEO	LVL7M	AUDIO + VIDEO

The project created allows to enable and disable the different stimuli when editing the levels, which allows for a more flexible approach to the design of the levels themselves. This way, it is possible to create new levels with different combinations of stimuli, or even change the stimuli used in a level, without having to change the logic of the game itself. This flexibility was created with the goal of allowing for future improvements or the creation of new levels, with different combinations of stimuli, without needing to create a new logic for each level individually, which would not only be time consuming, but also more prone to errors. This way not only avoids this, but also allows for future improvements and expansions of the project, as these can serve as a template for level creation.

3.5 Stimuli Implementation

This section will describe how both the haptic and auditory stimuli were developed and implemented into the game, as well as how they are controlled and triggered by the logic of the project.

In Section 3.4, we saw a general overview of how the different stimuli are implemented, but not how they are developed and integrated into the game. While the auditory stimulus is implemented through the use of audio clips, the process of selection and editing of these clips was not described. On the other hand, the haptic stimulus is implemented through the use of an external component, which was also not described in detail. The following sections will describe how these two stimuli were approached and implemented, including how the logic behind them works, and how the developed project controls them.

3.5.1 Haptic Stimuli Implementation

As mentioned before, the haptic stimulus used in the game is implemented through the use of an *Arduino Leonardo* board and vibration motors, in this case the *Lilypad Vibe Board*, which is directly connected to a vibration motor. Vibrations are commonly used as a form of haptic feedback in video games, but also in mobile devices, through push notifications, for example. In the context of this project, the vibrations are used not only as a form of feedback, but also as a way to directly communicate with the user, through the use of different patterns and intensities that can be attributed to different objects in the game.

Figure 3.6: Arduino Board and Vibration Motors Used in the Project

The control of the vibrations is done through the use of a script uploaded to the board, which is responsible for interpreting the commands sent by *Unity* and translating them into vibrations. This script also manages port communication problems, making sure the connection between the computer and the board is stable and workingcorrectly, which is important to avoid any possible issues that could arise during the experiment.

While *Unity* sends the commands and the *Lilypad Vibe Board* interprets them, the vibration motors are responsible for generating the vibrations themselves. Although these motors are capable of generating vibrations, they are not capable of generating widely different patterns and intensities, limiting the amount of different vibrations that can be used in the game. Even with these limitations, it was possible to create a set of vibrations that were distinct and recognizable between each other, therefore making it possible to use them as the main form of communication for the haptic stimulus. This limitation could have been avoided with a more accurate haptic device, one capable of generating a wider range of vibrations, as, for this study and motors used, we needed to manipulate the haptic signal by stringing together multiple haptic signals in order to create noticeable differences between each vibration, while a more accurate device could have circumvented this issue.

The communication and control done with the *Arduino Leonardo* board and the vibration motors is done through the use of a serial port, which allows *Unity* to send commands, as mentioned previously. These commands are sent in a specific format, in this case, a simple character string. As we saw in 3.4, the game is divided into two different tasks, each one using a grid of cards or squares, with each one containing a different object. Each object has a specific vibration pattern attributed to it, in this case, a specific character string. When the game logic triggers that object, be it either when the sequence is played or when the user selects it, the corresponding character string is sent to the *Arduino Leonardo* board, which, in turn, checks the string received and triggers the condition corresponding to that string, therefore generating the correct vibration pattern for the object selected. The same process is done when the user selects a wrong option, with the character string sent being replaced by a different one, in this case, a character that triggers the distinct vibration for a wrong selection. The vibration patterns only transmit the identity of the object, they do not transmit the location of the game tile itself.

While playing, one motor is attatched to the user's wrist, normally the non-dominant hand, so that they could freely interact with the touchscreen with their dominant hand, therefore avoiding any interference the motor cables could have during gameplay.

3.5.2 Audio Editing

On the auditory stimuli side, there was a need to find and edit all the different sounds used in this work. These sounds had to be easily distinguishable from each other, and also have a short duration, to avoid making the project too long, which could make the users not engaged or been able to memorize it. Using short sounds not only makes the experience more engaging, but it also allows for a better memorization of the sequence, as each sound is going to use less cognitive load from the user, therefore allowing them to focus on the sequence itself.

All the audios used were obtained from the *Freesound*[1] platform, which is a platform used to share and obtain sounds that are free to use without copyright attached. The sounds were chosen based on their duration, clarity, and distinctiveness. These factors all played an important role in the selection, as the sounds could not be too long, as that would make them way too prolonged, therefore making its editing process more difficult to compress into a shorter duration without losing the quality of the sound and without making it sound too artificial.

The editing process was carried out through the *Audacity*[2] program, a free and open source audio editor, which allowed us to edit the sounds as we needed, while maintaining or improving their quality. The editing process consisted in the following steps:

1. Cutting the audio clip to a maximum duration of one second, by only selecting the part of the sound that was the most clear and identifiable. For this, we selected the onset of the

[1] https://freesound.org/
[2] https://www.audacityteam.org/

sound, which is the part of the sound that is the most recognizable and distinct from the rest of the sound.

2. If the audio clip shows a big variation in its dynamic range, it is then compressed to reduce this variation.

3. Normalizing the audio clip to make sure every clip has the same volume.

4. Testing the audio clip to make sure it is clear and distinct from the other ones, and that it is still easily recognizable.

Once with the audio clips edited, they were then imported into *Unity*, where they were assigned to their respective objects in the game. This process was done in a similar way to the one seen in the haptic stimulus 3.5.1, where each object has a specific audio clip attributed to it, and when the game logic triggers that object, the corresponding audio clip is played. The same process is done when the user selects a wrong option, with the object's audio clip being replaced by a distinct one, only triggered for wrong selections.

The audios chosen are directly related to the objects in the game, with the animals, musical instruments and weather getting their respective sounds. The colors, on the other hand, do not have a specific sound related to them, so we used the visible spectrum of light to create a sound for each color, with the lowest frequency being attributed to he purple color, and the highest frequency being attributed to the orange color.

This color frequency to sound frequency mapping was done in a linear way, and followed the concepts of the study conducted by Miyamoto et al [17], where they conducted tests using two colors with two distinct sounds, and concluded that the human brain is able to associate colors with sounds, as long as the sound frequency is closely related to the color frequency.

3.6 Data Logging

After defining every variable we want to collect, we need to find a way to log the data dependent on our independent variables (in this case, the different stimulus and their combinations used and the minigame type). As such, a data logging system was implemented in the project, which is responsible for collecting and saving all the data generated during the experiment. This includes everything the participant does in the game, such as time taken per level and number of attempts for example.

This data logging system needs, first of all, to be able to save the data under a unique identifier, which is generated at the start of the game, and is used across all the different components of the project, such as the eye tracking and the questionnaire. The data pertaining to the experiment itself is saved in a JavaScript Object Notation (JSON) file, allowing us to structure the data in a readable and easy to understand way. On top of saving the participant ID, a procedure ID was also saved, which is used to identify the order of the levels played by the participant, with four different procedures being defined.

The JSON file saves the following data:

- Simon Says
  - Level Name
  - Task Completion Time
  - Error Count
  - Time when the level started and time when thelevel was closed
  - Completion Status
  - List of clicks per attempt, with timestamp of each click
- Memory Game
  - Level Name
  - Task Completion Time
  - Error Count
  - Time when the level started and time when the level was closed
  - Completion Status
  - List of clicks per card, with timestamp of each click.
  - List of pairs found and list of wrong pairs selected, with a timestamp for each pair.

The logging mechanism was implemented on a per level basis, meaning that the data for a specific level is only saved when the level is completed or closed by the time limit. Once every level is completed, the data is compiled into an unique JSON file, which is then saved in a specific folder containing all the data obtained from a given user during the experiment. In the next sections (Section 3.7 and Section 4.1.1) we will see how the data related to the eye tracking and the questionnaire is collected, and how it is integrated into the project.

3.7 Eye Tracking

The eye tracking is done with the use of a *Gazepoint GP3 HD* eye tracking device. This device is placed at the bottom of the touchscreen monitor, and is responsible for tracking and recording both the gaze behavior and the pupil dilation of the user, which will later be used as metrics to evaluate the cognitive load.

This device also uses its own software, *Gazepoint Control*, which is responsible for the calibration process, and *Gazepoint Analysis*, responsible for the recording and extraction of the data. The calibration is done right at the start of the experiment, and it is during this calibration that the software will record the initial pupil size of the user, which is used as a baseline for the pupil diameter changes we will be analyzing later. Each participant has different pupil sizes, therefore it is important to have this baseline done for every user, in order to make the results obtained uniform and comparable with each other.

The *Gazepoint Analysis* software is then used to analyze the user's behavior during the experiment, while also allowing to see if there are any issues with the eye tracking, such as loss of tracking. If these issues occur, the data obtained can be discarded, as it would not be accurate or useful and could negatively impact the study. While this software is capable of showing the heatmap of the user's gaze, and graphs of the pupil diameter changes, it also allows us to extract the raw data obtained in a *.csv* file, which will be the format used for this study, giving us more control over the analysis process.

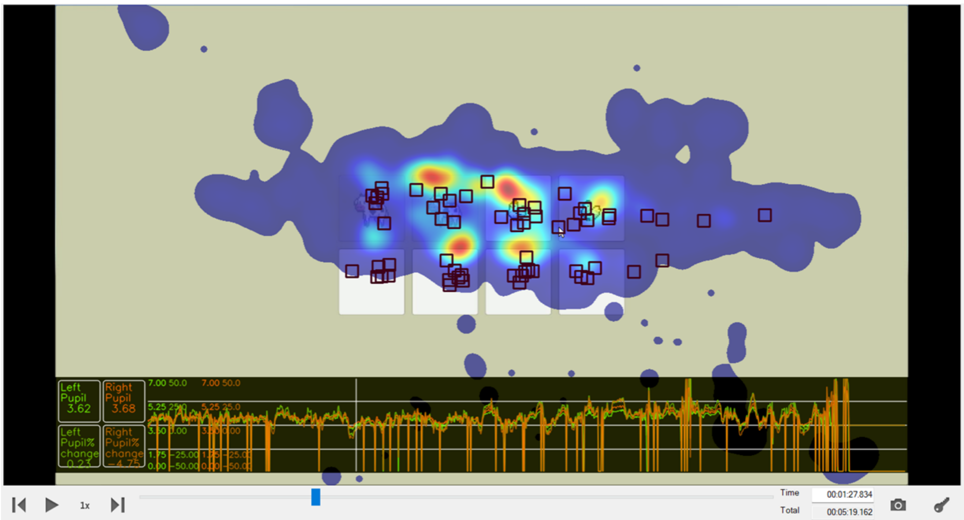


Figure 3.7: Example of the results obtained from the Gazepoint Analysis, showing the heatmap of the user's gaze and the graph of the pupil diameter changes

As shown in Figure 3.7, the software creates a heatmap of the user's gaze and shows the changes in pupil diameter during the experiment. Although this information is useful, it can be improved upon, as the current heatmap shows the gaze behavior from the start of the experiment until the point selected, and not the behavior for an individual level. The same goes for the pupillometry, which still has the blinks included (with the blinks being the sudden drops in the graph), and is harder to do an analysis per level. Due to these reasons, only the raw data extracted in the *.csv* file is used for this study.

3.8 Summary

The final experiment setup is based on a 2D game, divided into two different types of tasks, each one with its own levels, but keeping the use of the same stimuli. The two types of tasks are based on the "*Simon Says*" and "*Memory Game*" games, with each level containing one or more stimuli, be it either a visual, auditory or haptic stimulus, or a combination of two or more. These tasks were chosen because of their simplicity and minimalistic design, but also because of their ability to be easily understood, be it either from past experience or from the instructions given before a participant starts the experiment. These tasks are also different enough from each other, even if they are in the same memory related context, each one allows for a different approach to the study, therefore possibly showing different results from the same stimuli.

The implementation of the different devices used in the experiment, such as the eye tracker and the haptic device, was also explained, along with how they are integrated into the project itself. The eye tracker is used to collect metrics related to the cognitive load of the user, while the game itself collects the performance metrics. These metrics are then logged and saved for future analysis, in turn making it possible to evaluate the results obtained.

On the other hand, the haptic device is used to introduce vibrations into the game, which differently from regular implementations of haptics, who are used only as feedback, are used as a form of communication with the user, through the use of different patterns and intensities that can be attributed to different objects in the game. This way, we can evaluate if it is possible to use haptics as an unique stimulus in a memory related task, or if it has to be combined with other stimulus to have a positive effect on an user. Besides this stimulus, the auditory stimulus was also covered in this chapter, showing how the different sounds were obtained and edited.

The efforts seen in this chapter were all focused on the development of a project that could not only reach the goals of this study, but also be accessible to a wider range of participants. The decisions made during the development process, such as the use of a touchscreen monitor and the simplification of the original idea, were all done with this goal in mind, as we wanted to avoid limiting our research pool to a small and specific group of people, such as those familiar with gamepad controllers and 3D environments. The final project is a reflection of these efforts, and with the evaluation of the results obtained and participant feedback, we can see if these efforts were successful or not.

Chapter 4

Evaluation and Results

Having explored the methodology in Chapter 3, we can now move on to the evaluation of the proposed approach and analysis of the results. In this stage, it is important to recall the research questions defined in Section 1.3 of Chapter 1.

- **RQ1:** How do different types of visual, auditory and haptic stimuli affect cognitive load and player performance in multimodal memory games?
- **RQ2:** Is the impact on the cognitive load and player performance considerably different with and without combined sensory feedback?
- **RQ3:** How do these three stimuli interact with each other to influence the final results?

To get an answer to the aforementioned research questions, it was necessary to perform a number of tests with users in a controlled environment. The results obtained from these tests will support that there is, in fact, an impact on cognitive load and player performance when using different types of stimuli, independent if they are used alone or in a combination.

First this chapter will cover topics related to the evaluation process, such as the experimental setup and the order of the tasks presented. The questionnaire developed for this study is also presented, discussing every stage involved in it. After the evaluation, the chapter will present the results obtained from the tests, including the processing of the data collected and how it was analyzed. Finally, the results will be discussed and compared with the RQ defined in Section 1.3 of Chapter 1, connecting the results with the RQ directly in order to draw conclusions and answers to them.

4.1 Evaluation

To evaluate the results obtained, we first need to understand the structure of the project and how the different types of data were collected. As we saw in Chapter 3, the project collects data both from the game itself, and from the *Gazepoint GP3 HD* eye tracker. Besides this, it is also possible to collect data from other sources, such as the NASA-TLX questionnaire.

This section will explore the questionnaire created for this study, the structure it follows, and how the NASA-TLX questionnaire is integrated into it. After this, the experimental setup will be presented, showing the steps that were conducted during the experiment, and how each participant was guided through the different stages of the test.

4.1.1 Integration of the Questionnaire

The questionnaire used is divided into three different sections: a pre-experiment section, a intermediate section, and a post experiment section. The questionnaire was carried out through the *Google Forms* platform, which allowed for an user friendly and simple way for the participants to answer the questions contained in it, while also allowing for an easy extraction of each user's answers, through a *.csv* file to be later used in the analysis of the results.

The pre-experiment section is used to collect demographic data, such as age, how often they play video games, and familiarity with memory based games, such as, for example, jigsaw puzzles, spot the difference games, memory card/pairing games, and so on. This data is useful to understand the background of the participants, and it helps observe if there are any correlations between the players' background with the results obtained during the experiment. In here, the participants also give their informed consent to participate in the experiment, authorizing the collection of their data for the sole purpose of this research.

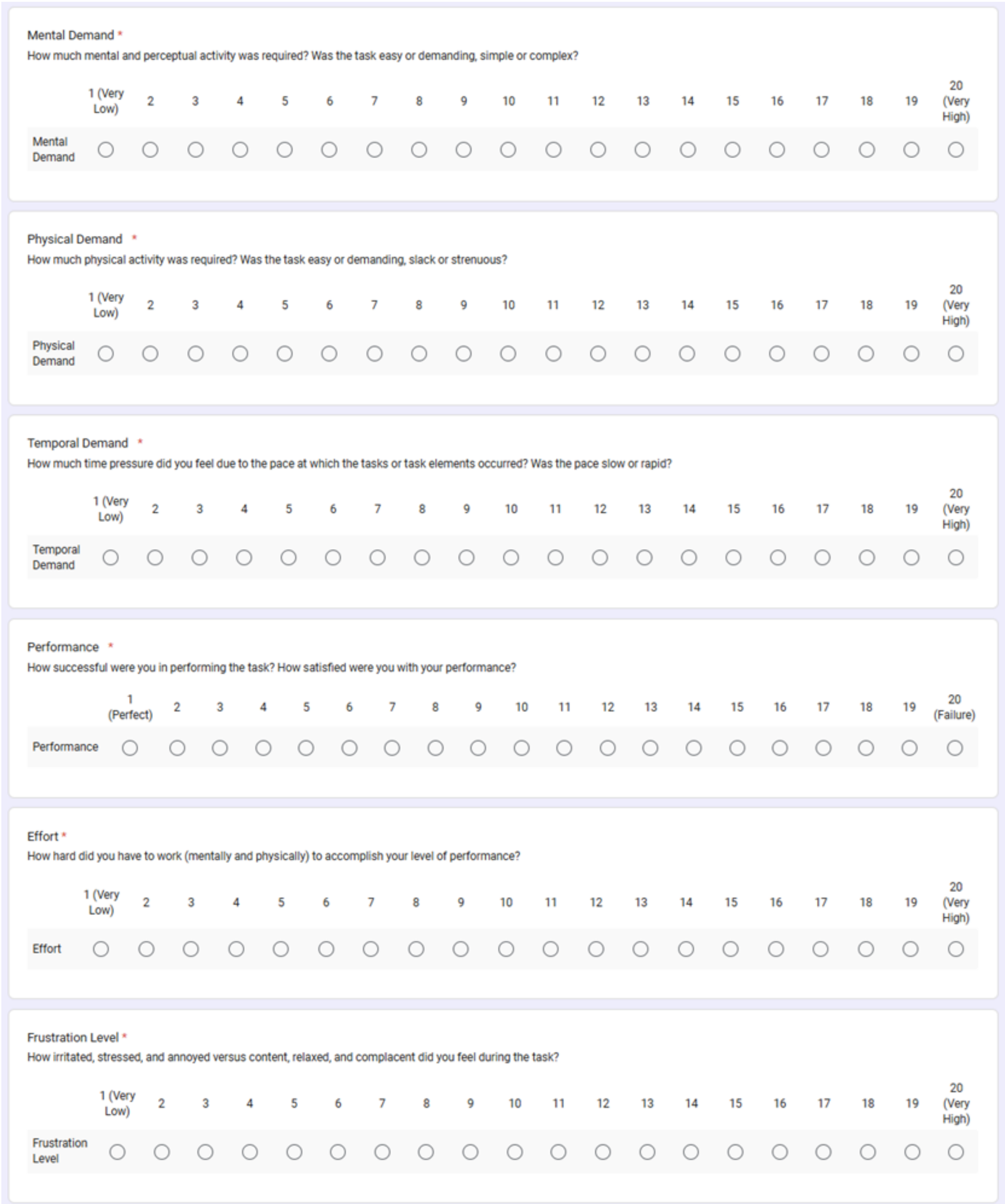
The intermediate section is divided into two different parts, one for each type of game, in this case the "*Simon Says*" and the "*Memory Game*". In this section, a NASA-TLX questionnaire is used to evaluate the cognitive load, adding another possible measure for this metric besides the gaze behavior and pupil dilation seen in 3.7. This section is filled out after the player completes every level of each game type, which allows the player to successfully evaluate their performance between the two games and give a more accurate evaluation in the questions being asked. In here, the player must also answer which levels were more complicated for them, and which ones were easier, making it possible to understand if a certain type of stimulus was more difficult to process than others, and if there is any correlation between difficulty and stimulus used.

Finally, the post-experiment section is used to collect general feedback and possible suggestions, such as future improvements that could be implemented in the game, or limitations found that could have impacted the experience. This section also includes both a question about which stimuli the player preferred, and which one they disliked the most, giving us a better understanding of how the participants feel about the different stimuli implemented in the game.

4.1.2 Experimental Setup

Now with all the different components of the project defined, explained and implemented, it is time to define the experimental setup, which will be used to conduct the experiment with the participants in a controlled environment. The setup is as follows:

- The user starts by filling out the pre questionnaire, which is used mainly to collect the user's informed consent and demographic data.



Mental Demand *
How much mental and perceptual activity was required? Was the task easy or demanding, simple or complex?

	1 (Very Low)	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20 (Very High)
Mental Demand	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○

Physical Demand *
How much physical activity was required? Was the task easy or demanding, slack or strenuous?

	1 (Very Low)	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20 (Very High)
Physical Demand	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○

Temporal Demand *
How much time pressure did you feel due to the pace at which the tasks or task elements occurred? Was the pace slow or rapid?

	1 (Very Low)	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20 (Very High)
Temporal Demand	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○

Performance *
How successful were you in performing the task? How satisfied were you with your performance?

	1 (Perfect)	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20 (Failure)
Performance	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○

Effort *
How hard did you have to work (mentally and physically) to accomplish your level of performance?

	1 (Very Low)	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20 (Very High)
Effort	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○

Frustration Level *
How irritated, stressed, and annoyed versus content, relaxed, and complacent did you feel during the task?

	1 (Very Low)	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20 (Very High)
Frustration Level	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○

Figure 4.1: Nasa-TLX Questionnaire Used in the Project

- Start with checking the auditory stimulus, then haptic stimulus, and finally the eye tracking. Ask the user if each component is working correctly or feels comfortable, and if not, try to solve the issue before starting the experiment. For the eye tracking, it is necessary to go through the calibration process, ensuring with the user that the tracking is working correctly.
- The user then starts the experiment, starting with one of the four different procedures defined. The user only plays one of the procedures, and starts either with the "*Simon Says*" or the "*Memory Game*", depending on the orders given to them.
- Once both minigames are completed, the user fills out the intermediate questionnaire, which includes the NASA-TLX questionnaire for each game individually, and the questions about which levels were more difficult and which were easier.
- Finally, the user fills out the post questionnaire, which includes the questions about which stimuli they preferred and which they disliked the most, as well as general feedback and suggestions for improvement.

This setup is designed to be simple and straightforward, while also avoiding any issues that could arise during the experiment, such as problems with the equipment used, or issues with the understanding of the game mechanics. During each step of the setup, the user is allowed to ask any questions they might have, as long as they do not interfere with the experiment itself, such as asking about the mechanics of the project while playing, for example. This way, it is possible to guarantee that the results from the user are as accurate as possible, while also making sure the user is comfortable and relaxed.

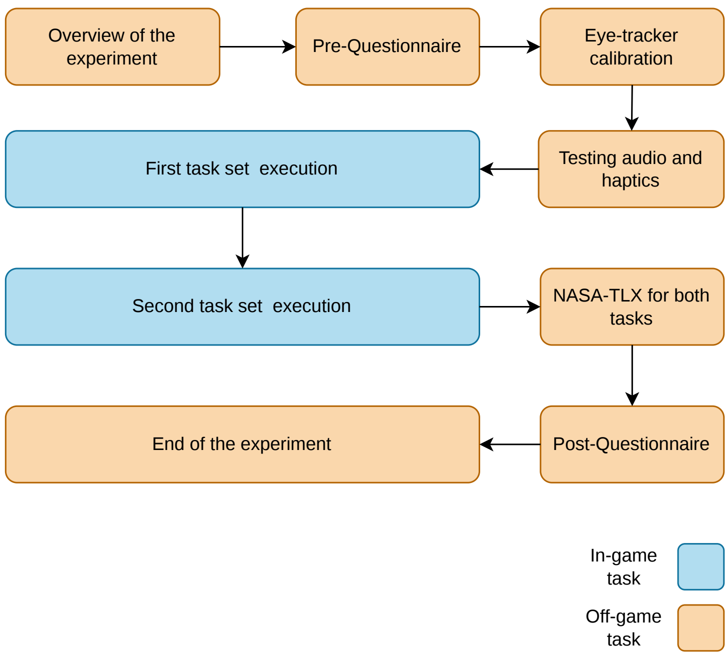


Figure 4.2: Diagram of the experimental flow

Figure 4.2 shows a diagram of the experimental flow described previously, showing the different steps that are taken during each participant experiment. These steps can be divided into two different phases, in-game and off-game, with the in-game phase being the one in which the participant is actively playing the game, and the off-game phase being the one in which the participant is either filling out the questionnaire or having the equipment checked/calibrated.

It is also important to mention that, every time a experiment is conducted, the light conditions of the room need to be the same. This is due to the eye tracker used being sensitive to the light conditions. For that reason, the room during the experiments was kept dimly lit, with the lights turned off and the curtains closed, in order to avoid major changes in the light conditions that could affect the data collected from the eye tracker.

The eye tracker used in this study, while being a good tool for measuring eye movements and pupil diameter, would lose the signal of the user's eyes if the user moved their head too quickly or if there were sudden changes in light conditions. This made it difficult to obtain accurate and consistent data, as this would happen even in a totally dark environment, with no outside light to interfere with the tracker. This meant that during the tests it became necessary to constantly monitor the eye tracker activity in the *Gazepoint* software in order to ensure that the data being collected was accurate, and in some cases, notify the user to adjust their position in order for the tracker to regain the correct eye position.

Still regarding the eye tracker, this device only worked correctly when on a specific distance and angle from the user's eyes. This meant that the user had to be positioned in a specific way, in most cases, standing up and looking down at the screen, as seen on Figure 4.3. This positioning was not ideal for tall users, as they would have to slightly bend their backs in order to be able to see the screen, but this was one of the only ways to ensure that the eye tracker would work with less interference. The eye tracker also did not work correctly with users that wore glasses, as the reflection of the glasses would interfere with the infrared light emitted by the eye tracker, making it impossible for users who can not see without glasses and who do not use contact lenses to participate in the study, therefore reducing the possible number of participants.

4.1.3 Order

As seen in Table 3.1, each level is designed to be played with a specific type of stimulus, or a combination of them. The order in which the levels are played is important, as it can have an impact on the results obtained, as mentioned previously.

For that reason, the order in which the levels are played is defined by four different procedures, which are as follows:

For simplicity purposes, the levels are numbered from 1 to 7, using only their identifier instead of the full name seen on the Table 3.1. The order of the levels is designed to avoid the possible bias that could be introduced by the order, as seen in Chapter 3. The four different procedures are designed to make it so the levels show up in the maximum number of different positions possible, with the levels showing both near the start and near the end of the order, depending on the procedure being used.

Table 4.1: Level Order By Procedure

Procedure	**Level Order Simon Says**	**Level Order Memory Game**
1	1-2-3-4-7-6-5	1-2-3-4-5-6-7
2	2-1-3-6-5-4-7	1-4-2-3-6-5-7
3	7-1-5-4-6-2-3	7-6-1-4-2-5-3
4	6-4-7-5-3-2-1	7-3-5-6-2-1-4

On top of this, the order in which the participant plays the two different tasks is also important, with participants with an odd identifier starting with the "*Simon Says*" task, while participants with an even identifier start with the "*Memory Game*" task. Each identifier is assigned based on the procedure being used, and the number of participants that have played that procedure, with the first participant to play the procedure 1 being assigned the identifier 11, the first participant to play the procedure 2 being assigned the identifier 21, and so on.

The procedures are then played in a round robin fashion, with the first two participants playing the procedure 1, one starting with the "*Simon Says*" task and the other starting with the "*Memory Game*" task, then the next two participants playing the procedure 2, and so on. This way, we ensure that each procedure is played by the same number of participants, without any procedure being played by a significantly higher number of participants than the others.

With these rules put in place it is possible to reduce the possible bias that could be introduced by the order of the levels, therefore making the results obtained more accurate.

4.2 Results

A total of 36 users participated in this study, one of whom is shown in Figure 4.3. The ages of these participants ranged from 10 to 65 years old, with an even distribution of both male and female participants. The participants were asked how often they play video games and how familiar they are with memory related games (such as jigsaw puzzles, spot the difference games, and the games used in this study), with the answers being input on a 5 point likert scale, with 1 being *Not Familiar at All* and 5 being *Very Familiar*, in the case of the familiarity question. For the frequency question, the options were *Never*, *Rarely*, *A few times per month*, *A few times per week* and *Daily*.

The Table 4.2 shows the distribution of responses to these questions.These distributions show that the participants of this study have a good variety of both frequency and familiarity with memory video games, which is important for a diverse and reliable number of results that cover all types of users.

Outliers were evaluated, with only one participant being considered an outlier, due to their young age (10 years old). Although their age was much lower than the rest of the participants, their results were not discarded, as they were still able to complete the experiment and showed legitimate attempts to do so.

Table 4.2: Participant Responses to Frequency and Familiarity Questions

Frequency of Use		**Familiarity (1–5)**	
Response	**Count**	**Score**	**Count**
Daily	10	5	12
A few times per week	5	4	9
A few times per month	7	3	9
Rarely	12	2	4
Never	2	1	2

The reporting of the results will use the following abbreviations, due to the number of times they will be mentioned and for simplicity purposes:

- Mean (M)
- Standard Deviation (SD)
- Simon Says (SS)
- Memory Game (MG)

Finally, we can also refer to the Dependent Variable (DV) of the experiment, which for the cognitive load is *pupil diameter changes*, *fixation duration* and *fixation count*, and the *NASA-TLX* results. For the player performance, the *error count* and the *TCT* will be used.

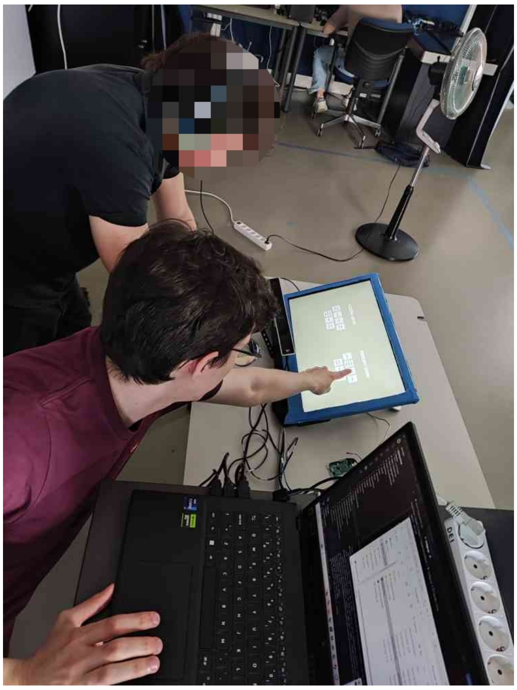


Figure 4.3: Participant in the Experiment

4.2.1 Preliminary Analysis

Before starting the evaluation of the results, it is important to perform a preliminary analysis. This analysis includes the cleaning of the data, as well as possible exclusions of participants due to the eye tracking data being too unreliable, which can happen when the eye tracker is not properly calibrated or lost track of the user's eyes systematically.

After an analysis of every participant's eye tracking data, it was possible to observe that four different participants had to be excluded from the analysis due to the aforementioned data. Although these participants' eye tracking data were not usable, their NASA-TLX and performance data was still included in the analysis, as these were valid and could still provide useful information for the evaluation.

A Shapiro-Wilk normality test was performed on all the DV per condition, in this case, the seven different stimuli conditions (visual, auditory, haptic, visual + auditory, visual + haptic, auditory + haptic and visual + auditory + haptic), for each type of minigame (SS and MG), giving us a total of 14 different conditions to test per DV. The results of the normality test showed that the normality was violated for at least one condition ($p < 0.05$). As normality could not be assumed under all conditions for any DV, non-parametric tests were used for the analysis of the results, in this case, the Friedman test, with a Kendall's W being applied to the results obtained. Finally, a post-hoc analysis with the Wilcoxon signed-rank test was performed to determine which conditions were significantly different from each other.

4.2.2 Player Performance

First, we will be looking into the level completion for each condition. While this can not be directly used as a performance metric, as it does not take into consideration how well the player performed but only if they were able to complete the level or not, it is still an important metric to observe how future results can be affected by the fact that some users were not able to complete the level.

Table 4.3 shows the level completion for both minigames. It is possible to observe that the SS minigame had a lower level completion rate overall, while the MG minigame had a 100% level completion rate for all conditions, which could possibly indicate that the SS minigame was more difficult than the MG minigame. It is also possible to observe that the conditions with the haptic stimulus had a lower level of completion rate, especially when the haptic stimulus was used alone, which could indicate that the haptic stimulus had a negative impact on the users, at least for the SS side.

To assess the impact of the different stimuli on player performance, we will be looking into a set of quantitative metrics, such as the error count and the TCT. These metrics will be analyzed for each of the seven different stimuli conditions, as well as for both minigames (SS and M), as mentioned previously. The Table 4.4 shows the M and SD of the error count for each condition, under both minigames. Meanwhile, the Table 4.5 shows the M and SD of the TCT for each condition.

Table 4.3: Level Completion By Minigame

Condition	Level Completion SS (%)	Level Completion MG (%)
Audio	71.9%	100%
Video	100%	100%
Haptic	3.1%	100%
Audio + Video	96.9%	100%
Audio + Haptic	81.2%	100%
Video + Haptic	96.9%	100%
Audio + Video + Haptic	100%	100%

Table 4.4: Error Count Metrics By Minigame

Condition	ErrorCount SS M (SD)	ErrorCount MG M (SD)
Audio	2.47 (2.33)	2.62 (1.36)
Video	0.97 (1.20)	2.50 (1.19)
Haptic	6.13 (1.99)	3.97 (2.91)
Audio + Video	0.59 (1.34)	2.09 (1.06)
Audio + Haptic	1.97 (2.35)	3.00 (2.06)
Video + Haptic	0.71 (1.42)	2.22 (1.10)
Audio + Video + Haptic	0.56 (1.04)	2.06 (1.11)
Friedman Test		
χ^2 (p)	91.41 (0.00)	14.03 (0.03)
Kendall's W		
W	0.476	0.073

It is possible to observe that the conditions with the lowest error count are the ones that include the visual stimulus, with the video being combined with one or two stimuli having an even lower error count. Although the visual stimulus shows good results, the auditory stimulus is also very promising, showing results that are close to the visual stimulus. On the other hand, the haptic stimulus shows a much higher error count, especially when used alone, but when combined, it appears to have a better performance. Meanwhile, the Friedman test shows that there is a significant difference between the conditions ($p < 0.05$) for both minigames.

On the TCT side, a similar pattern can be observed, with the visual and auditory stimulus showing lower TCT individually and combined for both minigames and the haptic stimulus showing the highest TCT, especially when used alone. Again, for this DV, the Friedman test shows that there is a significant difference between the conditions ($p < 0.05$).

Table 4.5: TCT Metrics By Minigame

Condition	TCT SS M (SD)	TCT MG M (SD)
Audio	17.99 (16.18)	26.56 (8.40)
Video	19.07 (1.20)	19.41 (4.65)
Haptic	59.93 (5.93)	27.28 (17.41)
Audio + Video	14.19 (9.84)	21.14 (7.15)
Audio + Haptic	17.36 (15.92)	26.11 (9.58)
Video + Haptic	13.97 (9.55)	20.14 (5.49)
Audio + Video + Haptic	15.65 (10.23)	19.56 (5.52)
Friedman Test		
χ^2 (p)	61.37 (0.00)	37.51 (0.00)
Kendall's W		
W	0.319	0.195

4.2.3 Cognitive Load

Now, for the cognitive load analysis, we will be looking into the pupil dilation (pupil diameter) changes, fixation duration and fixation count. The tables will follow the same structure as the ones used in Subsection 4.2.2, with the M and SD of each condition for both minigames, as well as the Friedman test results. Tables 4.6, 4.7 and 4.8 show the results obtained for each of these DV respectively.

Table 4.6: Pupil Dilation By Minigame

Condition	Pupil Dilation SS M (SD)	Pupil Dilation MG M (SD)
Audio	4.40 (1.05)	4.41 (1.03)
Video	4.32 (1.00)	4.31 (0.98)
Haptic	4.47 (0.92)	4.87 (2.10)
Audio + Video	4.49 (1.07)	4.49 (1.29)
Audio + Haptic	4.42 (1.01)	4.77 (1.98)
Video + Haptic	4.31 (1.04)	4.34 (0.97)
Audio + Video + Haptic	4.31 (0.94)	4.41 (1.01)
Friedman Test		
χ^2 (p)	23.94 (0.00)	34.30 (0.00)
Kendall's W		
W	0.125	0.170

As seen on Table 4.6, the pupil dilation changes show very similar results between all conditions, regardless of whether it is for the SS or MG minigame. Although these results are close to each other, the Friedman test tells us that there is a significant difference between the conditions ($p < 0.05$) for both cases. Finally, Kendall's W shows that the effect size is small for both minigames, which indicates that while the difference between the conditions is significant, the stimuli ended up having a limited effect on pupil dilation changes, which is consistent with the narrow range of results obtained for this DV.

Table 4.7: Fixation Duration By Minigame

Condition	Fixation Duration SS M (SD)	Fixation Duration MG M (SD)
Audio	0.38 (0.12)	0.45 (0.19)
Video	0.60 (0.25)	0.42 (0.15)
Haptic	0.39 (0.15)	0.48 (0.21)
Audio + Video	0.59 (0.26)	0.40 (0.13)
Audio + Haptic	0.45 (0.17)	0.45 (0.18)
Video + Haptic	0.52 (0.20)	0.44 (0.19)
Audio + Video + Haptic	0.51 (0.24)	0.46 (0.21)
Friedman Test		
χ^2 (p)	65.90 (0.00)	13.21 (0.04)
Kendall's W		
W	0.343	0.069

Table 4.8: Fixation Count By Minigame

Condition	Fixation Count SS M (SD)	Fixation Count MG M (SD)
Audio	78.31 (70.51)	49.97 (24.76)
Video	34.81 (32.85)	36.63 (27.39)
Haptic	102.16 (77.55)	63.38 (42.30)
Audio + Video	32.50 (58.14)	40.63 (22.09)
Audio + Haptic	60.81 (57.79)	43.69 (27.39)
Video + Haptic	29.03 (42.53)	38.88 (21.40)
Audio + Video + Haptic	32.56 (36.30)	34.75 (21.70)
Friedman Test		
χ^2 (p)	66.55 (0.00)	34.54 (0.00)
Kendall's W		
W	0.347	0.180

Now on Table 4.7 and Table 4.8, we can observe that levels that contain the visual stimulus have a longer fixation duration and a lower fixation count, which indicates that the users were able to focus on the relevant areas of the screen for a longer period of time when compared to the other conditions, which could be related to the need of lower cognitive load. Meanwhile, the haptic stimulus shows a lower fixation duration and a higher fixation count, which could be directly related to the users not being able to focus on an unique area of the screen, which in turn increases their cognitive load. The Friedman test shows that there is a significant difference between the conditions ($p < 0.05$) for both DV, while Kendall's W shows that the effect size is small only for MG, while for SS the effect size is medium, indicating that the stimulus had a more significant impact on the SS minigame than on the MG minigame.

4.2.4 Post-hoc Analysis

As mentioned before, the Friedman test is not enough to understand the differences between the conditions, as it only tells us that there is a significant difference between them, but not which conditions are different from each other. For this reason, a post-hoc analysis using the Wilcoxon Signed-Rank test is necessary, as it solves this problem by comparing each condition with each other. On top of this, a Bonferroni correction was applied to the results obtained, in order to reduce the chances of obtaining false positive results. These results are shown in Appendix B.

With the observing of the results pertaining to the SS tasks, it becomes clear that the condition with the haptic stimulus alone shows a significant difference with all the other conditions. This is consistent with every performance metric analyzed, with also some of the cognitive load metrics (fixation duration and fixation count), showing that the haptic stimulus had, more than likely, a negative impact on the users when used alone. Meanwhile, the auditory stimulus alone also shows differences with some of the conditions, although not as many as the haptic stimuli, it is still considerable enough to indicate that this stimulus is also not as effective as the visual stimulus, which shows the best results overall, with no significant differences with any of the other conditions. Meanwhile, the combination of the three stimuli also shows good results, showing that combining all the stimuli together can be a good strategy to achieve better results.

On the MG side, the results are not as clear as in SS, with both auditory and haptic stimulus alone showing significant differences with some of the other conditions. A clear difference between the two can not be observed, as the auditory stimulus shows significant differences for DV such as TCT and fixation count, while the haptic stimulus shows these differences for the error count and fixation count, mainly. The visual stimulus alone and combined shows good results, with it still being the preferred stimuli for the tasks. The combination of all three stimuli maintains the pattern observed in SS, with it showing good results and no significant differences with the majority of the other conditions.

These results mainly show that both auditory and haptic stimulus alone are not as effective as the visual stimulus, while the combination of them shows a positive impact on the users. These results will be further explored in Section 4.3, where we will compare the conclusions obtained from the results with the research questions defined in Section 1.3.

4.2.5 Questionnaire

In the questionnaire, created using the NASA-TLX template described in Section 4.1.1, all the users were asked to rate different categories on a scale of 1-20, which are: Mental Demand, Physical Demand, Temporal Demand, Performance, Effort and Frustration. The answers were completely subjective, as they were based on the user's perception while playing. The Table 4.9 shows the total scores for both SS and MG, after calculating the average.

Table 4.9: NASA-TLX TotalScores By Minigame

NASA-TLX TotalScores SS M	NASA-TLX TotalScores MG M
6.64	4.44

As seen on Table 4.9, SS had a higher total score than MG, which indicates that the users perceived the SS minigame as the most demanding one between the two. Although this happened, both tasks had a low total score, which indicates that the users did not perceive them as very demanding. While the majority of the users rated both tasks with a low score, some outliers were spotted, with some users having their total score for SS above a 10, meaning that they perceived the task as moderately demanding. Meanwhile, MG did not have such outliers, with the highest total score still being below a 10.

4.3 Discussion

To address the results obtained and the research questions previously defined, we will be further exploring and interpreting the results, using the conclusions and patterns observed in the Subsection 4.2.4 as a base for the discussion.

As we briefly saw, the haptic stimulus alone had a negative impact on the users, with the auditory stimulus alone also impacting the users, although not to the degree seen in the haptic stimulus. The visual stimulus, however, showed the best results overall, with it being the preferred stimulus to use, with barely any significant differences from the other conditions. With this in mind, we can obtain an answer to the first research question, **RQ1**: *How do different types of visual, auditory and haptic stimuli affect cognitive load and player performance in multimodal memory games?*.

For **RQ1**, the visual stimulus does not have a negative impact on the cognitive load, with no major differences being observed on the DV related to this field (pupil dilation, fixation duration and fixation count). The results show that there is not a significant difference between the visual stimulus and the other conditions, with the signficant differences being observed mainly on the haptic stimulus alone, and in some cases of the auditory stimulus alone. This indicates that for cognitive load, the visual stimulus is the most effective of the three, with it being able to provide the necessary information for the users to complete a task without increasing their cognitive load. This is also supported by the subject opinion of the users, where in the questionnaire they rated this stimulus as the least demanding of the three, while the haptic stimulus was rated the hardest one.

On the performance side, the visual stimulus also shows the best results, with both the error count and TCT being lower than the other conditions, with the haptic stimulus alone showing the worst results. Meanwhile, the auditory stimulus and haptic stimulus show a higher error count and TCT, with the haptic stimulus alone being the worst of the three. The literature seen on Chapter 2 supports this results pertaining to the haptic stimulus, as Jamalian et al. [13] indicated that the haptic stimulus did show an increase in the cognitive load of the users involved in that particular study, in turn worsening the results,which is consistent with what we observed in this study.

For **RQ2**: *Is the impact on the cognitive load and player performance considerably different with and without combined sensory feedback?*, we see that the combination of the stimuli does have a small positive impact on the users, be it either a combination of only two stimuli or all three. The results show that the visual stimulus is the dominant factor in both cognitive load and performance, with the combination of this stimulus with the auditory and/or haptic showing better results in some metrics. While these results were observed, the post-hoc analysis shows that there is no significant difference between the visual stimulus alone and its combinations (Video vs Video + Audio + Haptic had p=1.00, meaning no significant difference), with the results being very similar and improved slightly in some cases, with the same happening for the performance metrics. This indicates that a multimodal approach can be beneficial, but it is not a major improvement over the visual stimulus alone, with this approach being potentially more useful to improve the user experience and comfort, rather than to improve the results obtained by a big margin.

Finally, for **RQ3**: *How do these three stimuli interact with each other to influence the final results?*, the stimuli do appear to interact within each other, with mainly positive outcomes being observed. For example, the haptic stimulus alone creates a negative impact on the users, but when combined with another stimulus it shows that it can turn out to be beneficial. We saw that on the SS minigame, the haptic alone had a low level completion rating, close to 0%, but when combined with either auditory or visual, the level completion rating increased to 81.2% and 96.9%, respectively, and with the combination of all three stimuli, the level completion rating increased to 100%. This shows that haptic itself is not a negative impact to the users, it only becomes harmful when used alone, as it is unable to provide the necessary information for the users to complete their tasks. This happens mainly on the SS minigame, as it is more complex than the MG minigame, indicating that the complexity of the task itself can also have an impact on how the stimuli is perceived by the users, and if that stimuli is able to provide the necessary information for the users to complete their tasks.

Meanwhile, the visual stimulus and auditory stimulus alone may show better results than the haptic stimulus alone, but when they are combined with another stimulus, their results improve even further. The combination of all three stimuli shows the best results overall, be it either on the cognitive load or performance side, with the users being both able to complete the tasks with a lower cognitive load and a better performance, for both minigame types. This indicates that the stimulus can interact with each other to enhance their effectiveness, as studies such as [20] and [7] observed the same pattern, with the introduction of a second stimulus enhancing or even changing the perception of the first stimulus, therefore making it so that the users are able to process the information presented to them in different ways.

In conclusion, we were able to answer all the research questions initially proposed based on the results obtained. This study showed that different stimuli can have a different impact on the users, with some being more effective than others, and that the combination of the stimuli can be a good strategy to increase the performance and focus of the users, especially when the visual stimulus is included in that combination. The results obtained can therefore be considered promising, albeit there are limitations and future lines of work that can be explored, as identified in the next and final chapter.

Chapter 5

Conclusions

The study of cognitive load and player performance in videogames is a continuously evolving field, with new research and insights emerging regularly. As this area of study is still vastly underexplored, a motivation for this dissertation was to contribute to the understanding and development of this field.

In the related works, we saw the fundamental concepts of cognitive load, external stimuli and the evolution of videogames, from a tool of entertainment to a tool of education and cognitive assessment via Serious Games (SG). We also further explored all these concepts in the context of Human Computer Interaction (HCI), and how they can be used to improve the user experience by reviewing and analyzing the existing literature on the topic.

With the knowledge obtained from the literature review, we designed and developed a SG that properly integrates visual, auditory, and haptic stimulus, aiming to learn the impacts they have on cognitive load and player performance, either individually or in combination. The conceptual overview and architecture of the proposed solution was presented, subsequently showing how the implementation of the proposed solution was carried out.

In the implementation, we presented the first approach of the SG we thought of, and the issues that could arise if we further developed it. We then presented a second approach, which was the one that was implemented and used for the study. This second approach was then further explored, showing the different tasks involved and how the stimuli were implemented into them. The types of tasks (*Simon Says* and *Memory Game*) were chosen for their simplicity and flexibility, as they could be easily adapted to the different stimuli and their combinations. The implementation of the stimuli followed the same concept, with them being implemented through distinguishable cues that could be easily identified by the user.

The evaluation showed how the experimental setup would be conducted for each participant, including an explanation of the questionnaire that would be used in this stage. The results were obtained from a pool of 36 participants, making it so that we could reach the conclusion that the use of external stimuli can have a significant impact on both cognitive load and player performance, but it depends on the type of stimuli used and the specific task being performed. While using one stimulus may not be enough to significantly impact cognitive load and player performance, the combination of multiple stimuli can have a more pronounced effect. The use of external stimuli should be carefully considered, as it is necessary to first understand the specific task of the videogame and the desired outcome before deciding on which stimuli to use. The results themselves support this claim, as stimuli like haptic do not provide the user with enough information to complete the task. This, in turn, can lead to a higher cognitive load and lower player performance, therefore increasing the user's frustration and decreasing their overall experience.

In summary, we consider this study to be a valuable contribution to the field of cognitive load and player performance in videogames, providing insights and room for future research in this area. The findings of this study can be used to improve the design of videogames/SGs, making them more effective and engaging for players when using external stimuli like the ones explored in this study.

5.1 Limitations and Future Work

The study had some limitations that should be considered, such as the eye tracker and haptic device used. We previously mentioned that the eye tracker was not able to work with users that wore glasses, which reduced the number of possible participants that could take part in the study. We also mentioned that if an user was too tall, they would have to bend their back in order to be able to see the screen, which could provoke discomfort while playing. On the other hand, the haptic device used was not able to provide highly accurate haptics, making it so we had to manipulate the haptic signal in order to string together multiple haptic signals to create noticeable differences between each vibration.

While these limitations were pointed out in the document, there is some small limitations that can also be mentioned. One of them is the number of participants, as while 36 participants is a good number for a study, more participants would offer a bigger pool of data to analyze, therefore providing more accurate results. Another limitation is the variety of minigames used, as while the two minigames used were simple and flexible, and offered a good insight into the study, more minigames could have been used to provide a more in-depth analysis. With more time for both the development and gathering of participants, more minigames and more participants could have been used, but due to time constraints, we had to limit the number of minigames and reduce the experimental phase, in order to be able to finish the study in time. Future work in this project could explore these points, as well as solutions for the previously mentioned eye tracker and haptic device.

Future research in this area could explore the use of different types of external stimuli in different types of videogames that involve cognitive tasks, such as interactive storytelling games where the player decisions impact the story (*Until Dawn*[1], *The Oregon Trail*[2]) via audio or different haptics to point the player in the right direction, social deduction games (*Among Us*[3], *Town of*

[1]https://www.supermassivegames.com/games/until-dawn
[2]https://oregontrail.ws/games/the-oregon-trail/
[3]https://www.innersloth.com/games/among-us/

Salem[4]) via visual, auditory or haptic cues when investigating the environment, or even strategy games where the player needs to manage resources and their team equipment to win (*Baldur's Gate 3*[5], *Civilization*[6]), with strong vibrations or audios playing the player makes a wrong move. Future research could also explore the haptic stimuli individually, as while it may not be enough to give the user enough information to complete the task, there may be other contexts where it can have a positive impact on cognitive load and player performance.

On the study itself, some users reported that while they felt the difference in haptics, a better device could have made spotting that difference easier, supporting the limitation mentioned previously. Other users reported that the levels could have been longer, as they felt one minute per level was not enough. Others said adding an example level before the actual test would have been useful, as they felt that they needed to understand the game better before starting the test. Finally, some users reported that they would like to see more levels, so they could try different stimuli more than once, and see if they could perform better than before.

Overall, this study provides a foundation for future research, be it either in the same context of cognitive load and player performance, or in the project developed itself. While the limitations of the study may have affected the results obtained, the majority of the users reported that they enjoyed the experience, and that these limitations did not appear to have a significant impact on their performance and attention during the tests.

[4]https://www.townofsalem.io/
[5]https://baldursgate3.game/
[6]https://civilization.2k.com/

References

[1] Tobias Appel et al. "Predicting Cognitive Load in an Emergency Simulation Based on Behavioral and Physiological Measures". In: *2019 International Conference on Multimodal Interaction*. ICMI '19: INTERNATIONAL CONFERENCE ON MULTIMODAL INTERACTION. Suzhou China: ACM, Oct. 14, 2019, pp. 154–163. ISBN: 978-1-4503-6860-5. DOI: 10.1145/3340555.3353735. URL: https://dl.acm.org/doi/10.1145/3340555.3353735.

[2] Nils Ove Beese et al. "How Vibrotactile and Auditory Feedback Can Affect Performance in Search for Invisible Objects in Virtual Reality". In: *Proceedings of the European Conference on Cognitive Ergonomics 2024*. ECCE 2024: European Conference on Cognitive Ergonomics. Paris France: ACM, Oct. 8, 2024, pp. 1–6. ISBN: 979-8-4007-1824-3. DOI: 10.1145/3673805.3673808. URL: https://dl.acm.org/doi/10.1145/3673805.3673808.

[3] Margaret Bowen, Markel Benedicto Costa, and Andrea Lucia Reyes Burga. "Enhancing Balance Through Multisensory Gaming: Evaluating Haptic and Auditory Cues for Rehabilitation". In: *Proceedings of the 16th Biannual Conference of the Italian SIGCHI Chapter*. CHItaly 2025: CHItaly 2025: 16th Biannual Conference of the Italian SIGCHI Chapter. Salerno Italy: ACM, Oct. 6, 2025, pp. 1–7. ISBN: 979-8-4007-2102-1. DOI: 10.1145/3750069.3750301. URL: https://dl.acm.org/doi/10.1145/3750069.3750301.

[4] Federico Calà et al. "Pupil Dilation Dynamics in Serious Game Emotional Activation". In: *2024 IEEE Gaming, Entertainment, and Media Conference (GEM)*. 2024 IEEE Gaming, Entertainment, and Media Conference (GEM). June 2024, pp. 1–4. DOI: 10.1109/GEM61861.2024.10585786. URL: https://ieeexplore.ieee.org/document/10585786.

[5] Zhuang Chang et al. "The Impact of Virtual Agents' Multimodal Communication on Brain Activity and Cognitive Load in Virtual Reality". In: *Frontiers in Virtual Reality* 3 (Nov. 29, 2022). ISSN: 2673-4192. DOI: 10.3389/frvir.2022.995090. URL: https://www.frontiersin.org/journals/virtual-reality/articles/10.3389/frvir.2022.995090/full.

[6] Daniel Chin et al. "Interactive Rainbow Score: A Visual-centered Multimodal Flute Tutoring System". In: ().

[7] Sam Chin et al. "Purrfect Pitch: Exploring Pitch Interval Learning through an Audio-Haptic Interface". In: *Proceedings of the Augmented Humans International Conference 2025*. AHs 2025: The Augmented Humans International Conference. Masdar City, Abu Dhabi United Arab Emirates: ACM, Mar. 16, 2025, pp. 241–252. ISBN: 979-8-4007-1566-2. DOI: 10.1145/3745900.3746079. URL: https://dl.acm.org/doi/10.1145/3745900.3746079.

[8] Sam Chin et al. "Sound2Haptic: A Toolkit for Portable Multi-Channel Haptic Integration Across Multiple Form Factors and Devices". In: *Adjunct Proceedings of the 38th Annual ACM Symposium on User Interface Software and Technology*. UIST '25: The 38th Annual ACM Symposium on User Interface Software and Technology. Busan Republic of Korea: ACM, Sept. 28, 2025, pp. 1–4. ISBN: 979-8-4007-2036-9. DOI: 10.1145/3746058.3758993. URL: https://dl.acm.org/doi/10.1145/3746058.3758993.

[9] Wan Chong Choi, Wan Chong Choi, and Iek Chong Choi. "Investigating the Effect of the Serious Game CodeCombat on Cognitive Load in Python Programming Education". In: *2024 IEEE World Engineering Education Conference (EDUNINE)*. 2024 IEEE World Engineering Education Conference (EDUNINE). Mar. 2024, pp. 1–6. DOI: 10.1109/EDUNINE60625.2024.10500551. URL: https://ieeexplore.ieee.org/document/10500551.

[10] Peter Fritz et al. "Analysis of Visual Behavior and Cognitive Load of Pilots during Engine Failures with an Emergency Landing Assistance System in Mixed Reality". In: *Proceedings of the 18th ACM International Conference on PErvasive Technologies Related to Assistive Environments*. PETRA '25: The PErvasive Technologies Related to Assistive Environments. Corfu Island Greece: ACM, June 25, 2025, pp. 216–221. ISBN: 979-8-4007-1402-3. DOI: 10.1145/3733155.3733197. URL: https://dl.acm.org/doi/10.1145/3733155.3733197.

[11] Mustafa Can Gursesli et al. "Silent vs. Sound: The Impact of Uniform Auditory Stimuli on Eye Movements and Game Performance". In: *2025 IEEE Gaming, Entertainment, and Media Conference (GEM)*. 2025 IEEE Gaming, Entertainment, and Media Conference (GEM). July 2025, pp. 1–6. DOI: 10.1109/GEM66882.2025.11155778. URL: https://ieeexplore.ieee.org/document/11155778.

[12] Aline Hufschmitt, Stéphane Cardon, and Éric Jacopin. "Dynamic Manipulation of Player Performance with Music Tempo in Tetris". In: *26th International Conference on Intelligent User Interfaces*. IUI '21: 26th International Conference on Intelligent User Interfaces. College Station TX USA: ACM, Apr. 14, 2021, pp. 290–296. ISBN: 978-1-4503-8017-1. DOI: 10.1145/3397481.3450684. URL: https://dl.acm.org/doi/10.1145/3397481.3450684.

[13] Nima Jamalian et al. "The Effects of Hand Tracking on User Performance: An Experimental Study of an Object Selection Based Memory Game". In: *2022 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. 2022 IEEE International Symposium on Mixed and Augmented Reality (ISMAR). Oct. 2022, pp. 768–776. DOI: 10.1109/ISMAR55827.2022.00095. URL: https://ieeexplore.ieee.org/document/9995326.

[14] Petar Jerčič, Charlotte Sennersten, and Craig Lindley. "The Effect of Cognitive Load on Physiological Arousal in a Decision-Making Serious Game". In: *2017 9th International Conference on Virtual Worlds and Games for Serious Applications (VS-Games)*. 2017 9th International Conference on Virtual Worlds and Games for Serious Applications (VS-Games). Sept. 2017, pp. 153–156. DOI: 10.1109/VS-GAMES.2017.8056587. URL: https://ieeexplore.ieee.org/document/8056587.

[15] Xiaofu Jin et al. "Decoding Cognitive Load: Eye-Tracking Insights into Working Memory and Visual Attention". In: *Proceedings of the 2025 Symposium on Eye Tracking Research and Applications*. ETRA '25: 2025 Symposium on Eye Tracking Research and Applications. Tokyo Japan: ACM, May 26, 2025, pp. 1–7. ISBN: 979-8-4007-1487-0. DOI:

10.1145/3715669.3725864. URL: https://dl.acm.org/doi/10.1145/3715669.3725864.

[16] Srikoti Rishi Kumar et al. “A Novel Knowledge Semantic Video Game Prototype to Treat Patients with Absence Epilepsy”. In: *2023 5th International Conference on Advances in Computing, Communication Control and Networking (ICAC3N)*. 2023 5th International Conference on Advances in Computing, Communication Control and Networking (ICAC3N). Dec. 2023, pp. 1085–1091. DOI: 10.1109/ICAC3N60023.2023.10541340. URL: https://ieeexplore.ieee.org/document/10541340.

[17] Kenta Miyamoto et al. “Congruency of Color–Sound Crossmodal Correspondence Interacts with Color and Sound Discrimination Depending on Color Category”. In: vol. 14. 4. SAGE Publications, July 1, 2023, p. 20416695231196835. DOI: 10.1177/20416695231196835. URL: https://doi.org/10.1177/20416695231196835.

[18] Helio C. Silva Neto et al. “Cow Milking Game: Evaluating a Serious Game for Cognitive Stimulation with an Elderly Population”. In: *Proceedings of the International Symposium on Interactive Technology and Ageing Populations*. ITAP ’16. New York, NY, USA: Association for Computing Machinery, Oct. 20, 2016, pp. 44–53. ISBN: 978-1-4503-4746-4. DOI: 10.1145/2996267.2996272. URL: https://dl.acm.org/doi/10.1145/2996267.2996272.

[19] Myque Ouellette, Lauren Breeding, and Corey Clark. “Using Applied Cognitive Load Theory and Difficulty Analysis for Educational Game Design for Understanding and Transference of Literacy Skills in Adults”. In: *Proceedings of the 14th International Conference on the Foundations of Digital Games*. FDG ’19. New York, NY, USA: Association for Computing Machinery, Aug. 26, 2019, pp. 1–11. ISBN: 978-1-4503-7217-6. DOI: 10.1145/3337722.3337725. URL: https://dl.acm.org/doi/10.1145/3337722.3337725.

[20] Jamal R Paden, Arpit Narechania, and Alex Endert. “BiasBuzz: Combining Visual Guidance with Haptic Feedback to Increase Awareness of Analytic Behavior during Visual Data Analysis”. In: *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*. CHI ’24: CHI Conference on Human Factors in Computing Systems. Honolulu HI USA: ACM, May 11, 2024, pp. 1–7. ISBN: 979-8-4007-0331-7. DOI: 10.1145/3613905.3651064. URL: https://dl.acm.org/doi/10.1145/3613905.3651064.

[21] Stella Paschalidou. “Technology-Mediated Haptic Interactions in Dhrupad Vocal Music Pedagogy: What Next?” In: ().

[22] Yijun Qian et al. “Integrated Aerobic Exercise into Adult Second Language Learning in Virtual Reality Game”. In: *2023 IEEE Conference on Games (CoG)*. 2023 IEEE Conference on Games (CoG). Aug. 2023, pp. 1–4. DOI: 10.1109/CoG57401.2023.10333177. URL: https://ieeexplore.ieee.org/document/10333177.

[23] Emma Sophie Reichert et al. “Day ’n’ Nite: Investigating the Effects of Day and Night Time in the Real World and Game Worlds on Cognitive Performance”. In: *Companion Proceedings of the 2024 Annual Symposium on Computer-Human Interaction in Play*. CHI PLAY ’24: The Annual Symposium on Computer-Human Interaction in Play. Tampere Finland: ACM, Oct. 14, 2024, pp. 235–241. ISBN: 979-8-4007-0692-9. DOI: 10.1145/3665463.3678796. URL: https://dl.acm.org/doi/10.1145/3665463.3678796.

[24] Hannah J. Stewart et al. “Auditory Cognition and Perception of Action Video Game Players”. In: *Scientific Reports* 10.1 (Sept. 1, 2020), p. 14410. ISSN: 2045-2322. DOI: 10.1038/s41598-020-71235-z. URL: https://www.nature.com/articles/s41598-020-71235-z.

[25] Srikkanth Sundaram and Madeleine Steeds. “Frustration, Engagement, and Adaptation: Analyzing Player Emotions and Behaviors in Ori And The Blind Forest”. In: *Companion Proceedings of the Annual Symposium on Computer-Human Interaction in Play*. CHI PLAY ’25: The Annual Symposium on Computer-Human Interaction in Play. Pittsburgh USA: ACM, Oct. 13, 2025, pp. 88–95. ISBN: 979-8-4007-2023-9. DOI: 10.1145/3744736.3749340. URL: https://dl.acm.org/doi/10.1145/3744736.3749340.

[26] Agnieszka Szarkowska et al. “Watching Subtitled Videos with the Sound off Affects Viewers’ Comprehension, Cognitive Load, Immersion, Enjoyment, and Gaze Patterns: A Mixed-Methods Eye-Tracking Study”. In: *PLOS ONE* 19.10 (Oct. 7, 2024), e0306251. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0306251. URL: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0306251.

[27] Jacqueline Urakami, You Zhi Hu, and Mark Chignell. “Monitoring Cognitive Performance with a Serious Game: A Longitudinal Case Study on Online Cognitive Assessment Using Serious Games”. In: *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems*. CHI ’21: CHI Conference on Human Factors in Computing Systems. Yokohama Japan: ACM, May 8, 2021, pp. 1–7. ISBN: 978-1-4503-8095-9. DOI: 10.1145/3411763.3443431. URL: https://dl.acm.org/doi/10.1145/3411763.3443431.

[28] Chris Zimmerer et al. “Reducing the Cognitive Load of Playing a Digital Tabletop Game with a Multimodal Interface”. In: *CHI Conference on Human Factors in Computing Systems*. CHI ’22: CHI Conference on Human Factors in Computing Systems. New Orleans LA USA: ACM, Apr. 29, 2022, pp. 1–13. ISBN: 978-1-4503-9157-3. DOI: 10.1145/3491102.3502062. URL: https://dl.acm.org/doi/10.1145/3491102.3502062.

Appendix A

Literature Review

In order to better contextualize the areas of study involved in this dissertation, a set of relevant concepts are presented in this section. This Appendix serves as a supplementary material to the main text, providing additional details and context for the literature review, such as search queries and the inclusion and exclusion criteria.

To perform this analysis, a search of scientific articles related to the themes presented previously was put in place. This search took in consideration a vast number of open repositories and developed queries in order to filter the irrelevant information from the articles found. After that, an inclusion and exclusion process was put in place to further filter the articles obtained from the search queries.

A.1 Search Queries

As seen on 2.1, the articles obtained for the literature review were obtained through various repositories and databases. To be able to obtain the most relevant articles from these search engines, a set of search queries were developed.

The search queries used were refined around specific terms in order to cover the most amount of articles possible without cluttering that could show up from unrelated articles with similar search terms. These queries were formulated with the boolean terms (AND, OR) in order to combine important terms of this research together. These terms range from HCI to themes more intrisic to the study, like a variety of different approaches to the "*videogame*" keyword and all external stimuli previously mentioned (*audio*, *video* and *haptic*).

The following search queries were created for this study:

While only using two queries might make it look like the number of results obtained would not be ideal to collect enough information, the opposite happened. A vast number of results were obtained, with some repositories going above one thousand results. To avoid going through all of these articles some filters were applied. These filters included only searching for articles published on the last five years, as well as changing the order of the results to show the most cited articles first, if possible.

Table A.1: Search Queries

SQ ID	Search Query
SQ1	("VIDEOGAME" OR "VIDEO GAME" OR "GAMING" OR "SERIOUS GAME" OR "MULTIMODAL GAME" OR "MEMORY GAME" OR "MULTIMODAL MEMORY GAMES") AND ("COGNITIVE LOAD" OR "COGNITIVE HEALTH" OR "COGNITIVE PERFORMANCE")
SQ2	("VIDEOGAME" OR "VIDEO GAME" OR "GAMING" OR "SERIOUS GAME" OR "MULTIMODAL GAME" OR "MEMORY GAME" OR "MULTIMODAL MEMORY GAMES") AND ("COGNITIVE PERFORMANCE" OR "PLAYER PERFORMANCE" OR "COGNITIVE LOAD") AND ("AUDIO" OR "AUDITORY" OR "VISUAL" OR "VIDEO" OR "HAPTIC")

The focus of the first query (SQ1) was to find a direct correlation between videogames and cognitive load, with or without the need of external stimuli like audio or haptic for example. This query aimed to get a general view of how videogames can directly impact cognitive load, and what are the main factors that could influence the relationship between these two topics.

The second query (SQ2) was then created to introduce the external factors. With these factors we can get a insight look into what these stimuli may do to the cognitive load of a user while they are playing. This not only offers a more detailed insight to what can and can not affect cognitive load but it also helps support the basis of this study.

Overall, these two queries provided a good amount of articles that covered a variety of topics related to the themes of this study. With these articles obtained, it was then possible to start the inclusion and exclusion process of the literature review.

A.2 Inclusion and Exclusion Criteria

The inclusion and exclusion process used in this study followed a multi step approach. This approach started with the inclusion of all the articles that seemed adequate to this study, being closely followed by a stricter exclusion process to filter out the irrelevant articles obtained.

Before starting going in-depth into these processses, it is important to state how the selection criteria was defined. The selection criteria was based on the relevance of the articles to the research questions previously defined in Section 1.3 of Chapter 1. For that reason, articles covering topics such as in the cognitive load branch, like cognitive load theory and how to measure it, were considered relevant. Articles that further explored the relationship between videogames and cognitive load were also considered relevant, as well as articles that explores the impact of external stimuli on each one of these topics individually or combined.

A.2.1 Inclusion

For the inclusion process, we started by getting all the results obtained from the search queries previously introduced in A.1. Due to the high number of results obtained, a method was introduced to evaluate which articles would be included for further analysis.

The method was to read the title and abstract of each article, and save the articles that seemed relevant from those that did not. While doing this, the conference score and number of citations of each article were also taken into consideration, as these are good indicators of the quality of the article. Articles that scored above a B grade on the CORE conference ranking system were considered for inclusion, as well as articles with a moderate or high number of citations for their age. Articles whose conferences were not ranked on the CORE system were also considered for inclusion, but only after a careful analysis of their content and conference they were published in.

A.2.2 Exclusion

Contrary to the inclusion process, the exclusion process had a more strict and convoluted approach. This process started right after the inclusion process was completed, with a list of articles that were considered relevant for further analysis.

The exclusion process started with removing any duplicates found on the list containing all the articles selected for inclusion. This duplicate removal was necessary due to the fact that some articles were found on multiple repositories or showed up more than once due to an altered version of the same article being published on a different year. In these cases, only the most recent version of the article was kept for further reading.

After the duplicate removal was completed, an in-depth reading of these articles was conducted, taking notes and highlighting the most important parts of each of article. The reading of these articles went further than the abstract, as expected, but it followed a set order. First the introduction and conclusion were read, in order to get a better understanding of the article and its relevance to this study. If the article seemed relevant after this, the rest of the article was read, starting on the methodology and then going through the results and discussion. This reading process allowed to reduced the number of articles by a considerable amount, as some articles that seemed relevant by their title and abstract were not as relevant as expected when reading their results or methodology.

Also important to mention is that during this reading process, if an article referenced another article that seemed especially relevant to this study, that referenced article would also be analysed and potentially included in the final list of articles selected. This small step allowed to include some articles that would not have been found during the initial search queries, be it either by mistake or accidently omitted by the query itself.

With the reading process completed, the final number of articles selected was obtained. These articles where then used to obtain the necessary information to cover the topics to be presented in the the next chapter of this dissertation, the State of the Art.

A.3 Summary

At the end of this chapter, the final number of articles selected was 51. These articles were obtained through the search queries and the inclusion/exclusion process previously described. These articles end up covering a wide range of topics related to the themes of the study, allowing to get a better understanding of the current scene regarding cognitive load and the impact videogames and external stimuli have on it.

While we had a good number of articles to use, not all of them were used in the final dissertation, with 28 articles being cited in the final version. The articles that were not cited were either not relevant enough to the topics covered in the dissertation, either due to not containing enough information to support the topics covered, or because the information they contained was better covered by other articles collected. This does not mean that these articles were not relevant to the study, it just indicates that the information they contained was better covered elsewhere.

Appendix B

Post-hoc Wilcoxon Signed-Rank Test Results

Only pairs with $p < .05$ are shown. W = Wilcoxon test statistic; p = Bonferroni-corrected p-value. The significant results are highlighted with an asterisk (*).

Table B.1: Post-hoc Wilcoxon Results — Pupil Diameter (Simon Says)

Condition A	**Condition B**	**W(p)**	**Significant**
Audio	Audio+Haptic	126.0(1.000)	
Audio	Audio+Video	123.0(1.000)	
Audio	Audio+Video+Haptic	143.0(1.000)	
Audio	Haptic	159.0(1.000)	
Audio	Video	146.0(1.000)	
Audio	Video+Haptic	135.0(1.000)	
Audio+Haptic	Audio+Video	172.0(1.000)	
Audio+Haptic	Audio+Video+Haptic	75.0(0.106)	
Audio+Haptic	Haptic	116.0(1.000)	
Audio+Haptic	Video	66.0(0.047)	*
Audio+Haptic	Video+Haptic	78.0(0.136)	
Audio+Video	Audio+Video+Haptic	75.0(0.106)	
Audio+Video	Haptic	111.0(1.000)	
Audio+Video	Video	88.0(0.294)	
Audio+Video	Video+Haptic	65.0(0.043)	*
Audio+Video+Haptic	Haptic	133.0(1.000)	
Audio+Video+Haptic	Video	182.0(1.000)	
Audio+Video+Haptic	Video+Haptic	154.0(1.000)	
Haptic	Video	149.0(1.000)	
Haptic	Video+Haptic	115.0(1.000)	
Video	Video+Haptic	161.0(1.000)	

Table B.2: Post-hoc Wilcoxon Results — Fixation Duration (Simon Says)

Condition A	**Condition B**	**W(p)**	**Significant**
Audio	Audio+Haptic	91.0 (0.365)	
Audio	Audio+Video	10.0 (< .001)	*
Audio	Audio+Video+Haptic	66.0 (0.047)	*
Audio	Haptic	179.0 (1.000)	
Audio	Video	0.0 (< .001)	*
Audio	Video+Haptic	15.0 (< .001)	*
Audio+Haptic	Audio+Video	32.0 (0.001)	*
Audio+Haptic	Audio+Video+Haptic	158.0 (1.000)	
Audio+Haptic	Haptic	111.0 (1.000)	
Audio+Haptic	Video	42.0 (0.003)	*
Audio+Haptic	Video+Haptic	73.0 (0.089)	
Audio+Video	Audio+Video+Haptic	103.0 (0.809)	
Audio+Video	Haptic	14.0 (< .001)	*
Audio+Video	Video	175.0 (1.000)	
Audio+Video	Video+Haptic	100.0 (0.670)	
Audio+Video+Haptic	Haptic	97.0 (0.551)	
Audio+Video+Haptic	Video	81.0 (0.173)	
Audio+Video+Haptic	Video+Haptic	115.0 (1.000)	
Haptic	Video	5.0 (< .001)	*
Haptic	Video+Haptic	31.0 (0.001)	*
Video	Video+Haptic	118.0 (1.000)	

Table B.3: Post-hoc Wilcoxon Results — Fixation Count (Simon Says)

Condition A	**Condition B**	**W(p)**	**Significant**
Audio	Audio+Haptic	145.0 (1.000)	
Audio	Audio+Video	27.5 ($<$ .001)	*
Audio	Audio+Video+Haptic	56.5 (0.004)	*
Audio	Haptic	133.5 (0.876)	
Audio	Video	58.5 (0.004)	*
Audio	Video+Haptic	14.5 ($<$ .001)	*
Audio+Haptic	Audio+Video	72.5 (0.021)	*
Audio+Haptic	Audio+Video+Haptic	96.0 (0.180)	
Audio+Haptic	Haptic	87.0 (0.034)	*
Audio+Haptic	Video	102.5 (0.091)	
Audio+Haptic	Video+Haptic	20.5 ($<$ .001)	*
Audio+Video	Audio+Video+Haptic	192.5 (1.000)	
Audio+Video	Haptic	24.0 ($<$ .001)	*
Audio+Video	Video	158.5 (1.000)	
Audio+Video	Video+Haptic	206.0 (1.000)	
Audio+Video+Haptic	Haptic	40.0 (0.001)	*
Audio+Video+Haptic	Video	247.5 (1.000)	
Audio+Video+Haptic	Video+Haptic	179.5 (1.000)	
Haptic	Video	22.0 ($<$ .001)	*
Haptic	Video+Haptic	25.5 ($<$ .001)	*
Video	Video+Haptic	161.5 (1.000)	

Table B.4: Post-hoc Wilcoxon Results — Error Count (Simon Says)

Condition A	**Condition B**	**W(p)**	**Significant**
Audio	Audio+Haptic	110.0 (1.000)	
Audio	Audio+Video	22.0 (0.008)	*
Audio	Audio+Video+Haptic	20.0 (0.006)	*
Audio	Haptic	22.0 (< .001)	*
Audio	Video	32.0 (0.043)	*
Audio	Video+Haptic	22.0 (0.005)	*
Audio+Haptic	Audio+Video	17.5 (0.061)	
Audio+Haptic	Audio+Video+Haptic	33.0 (0.046)	*
Audio+Haptic	Haptic	21.5 (< .001)	*
Audio+Haptic	Video	76.0 (1.000)	
Audio+Haptic	Video+Haptic	12.0 (0.005)	*
Audio+Video	Audio+Video+Haptic	37.0 (1.000)	
Audio+Video	Haptic	6.5 (< .001)	*
Audio+Video	Video	64.0 (1.000)	
Audio+Video	Video+Haptic	61.0 (1.000)	
Audio+Video+Haptic	Haptic	0.0 (< .001)	*
Audio+Video+Haptic	Video	47.0 (1.000)	
Audio+Video+Haptic	Video+Haptic	73.0 (1.000)	
Haptic	Video	2.5 (< .001)	*
Haptic	Video+Haptic	3.5 (< .001)	*
Video	Video+Haptic	69.5 (1.000)	

Table B.5: Post-hoc Wilcoxon Results — Task Completion Time (Simon Says)

Condition A	**Condition B**	**W(p)**	**Significant**
Audio	Audio+Haptic	215.0 (1.000)	
Audio	Audio+Video	207.0 (1.000)	
Audio	Audio+Video+Haptic	222.0 (1.000)	
Audio	Haptic	18.0 (0.003)	*
Audio	Video	251.0 (1.000)	
Audio	Video+Haptic	187.0 (1.000)	
Audio+Haptic	Audio+Video	187.0 (1.000)	
Audio+Haptic	Audio+Video+Haptic	241.0 (1.000)	
Audio+Haptic	Haptic	23.0 (0.001)	*
Audio+Haptic	Video	228.0 (1.000)	
Audio+Haptic	Video+Haptic	160.0 (1.000)	
Audio+Video	Audio+Video+Haptic	237.0 (1.000)	
Audio+Video	Haptic	27.0 ($< .001$)	*
Audio+Video	Video	155.0 (0.868)	
Audio+Video	Video+Haptic	251.0 (1.000)	
Audio+Video+Haptic	Haptic	29.0 ($< .001$)	*
Audio+Video+Haptic	Video	215.0 (1.000)	
Audio+Video+Haptic	Video+Haptic	189.0 (1.000)	
Haptic	Video	3.0 ($< .001$)	*
Haptic	Video+Haptic	6.0 ($< .001$)	*
Video	Video+Haptic	132.0 (0.263)	

Table B.6: Post-hoc Wilcoxon Results — Pupil Diameter (Memory Game)

Condition A	Condition B	W(p)	Significant
Audio	Audio+Haptic	177.0 (1.000)	
Audio	Audio+Video	184.0 (1.000)	
Audio	Audio+Video+Haptic	204.0 (1.000)	
Audio	Haptic	122.0 (0.464)	
Audio	Video	112.0 (0.253)	
Audio	Video+Haptic	150.0 (1.000)	
Audio+Haptic	Audio+Video	106.0 (0.171)	
Audio+Haptic	Audio+Video+Haptic	120.0 (0.413)	
Audio+Haptic	Haptic	228.0 (1.000)	
Audio+Haptic	Video	8.0 ($<$ .001)	*
Audio+Haptic	Video+Haptic	57.0 (0.003)	*
Audio+Video	Audio+Video+Haptic	217.0 (1.000)	
Audio+Video	Haptic	118.0 (0.367)	
Audio+Video	Video	132.0 (0.807)	
Audio+Video	Video+Haptic	146.0 (1.000)	
Audio+Video+Haptic	Haptic	124.0 (0.521)	
Audio+Video+Haptic	Video	65.0 (0.006)	*
Audio+Video+Haptic	Video+Haptic	117.0 (0.345)	
Haptic	Video	43.0 ($<$ .001)	*
Haptic	Video+Haptic	84.0 (0.033)	*
Video	Video+Haptic	227.0 (1.000)	

Table B.7: Post-hoc Wilcoxon Results — Fixation Duration (Memory Game)

Condition A	Condition B	W(p)	Significant
Audio	Audio+Haptic	192.0 (1.000)	
Audio	Audio+Video	149.0 (1.000)	
Audio	Audio+Video+Haptic	208.0 (1.000)	
Audio	Haptic	165.0 (1.000)	
Audio	Video	160.0 (1.000)	
Audio	Video+Haptic	195.0 (1.000)	
Audio+Haptic	Audio+Video	134.0 (1.000)	
Audio+Haptic	Audio+Video+Haptic	206.0 (1.000)	
Audio+Haptic	Haptic	155.0 (1.000)	
Audio+Haptic	Video	170.0 (1.000)	
Audio+Haptic	Video+Haptic	170.0 (1.000)	
Audio+Video	Audio+Video+Haptic	144.0 (1.000)	
Audio+Video	Haptic	102.0 (0.238)	
Audio+Video	Video	177.0 (1.000)	
Audio+Video	Video+Haptic	197.0 (1.000)	
Audio+Video+Haptic	Haptic	194.0 (1.000)	
Audio+Video+Haptic	Video	153.0 (1.000)	
Audio+Video+Haptic	Video+Haptic	195.0 (1.000)	
Haptic	Video	110.0 (0.400)	
Haptic	Video+Haptic	140.0 (1.000)	
Video	Video+Haptic	204.0 (1.000)	

Table B.8: Post-hoc Wilcoxon Results — Fixation Count (Memory Game)

Condition A	Condition B	W(p)	Significant
Audio	Audio+Haptic	175.5 (1.000)	
Audio	Audio+Video	119.0 (0.410)	
Audio	Audio+Video+Haptic	42.5 (0.002)	*
Audio	Haptic	155.5 (1.000)	
Audio	Video	74.5 (0.024)	*
Audio	Video+Haptic	106.0 (0.194)	
Audio+Haptic	Audio+Video	158.5 (1.000)	
Audio+Haptic	Audio+Video+Haptic	131.5 (0.470)	
Audio+Haptic	Haptic	105.5 (0.189)	
Audio+Haptic	Video	190.0 (1.000)	
Audio+Haptic	Video+Haptic	193.5 (1.000)	
Audio+Video	Audio+Video+Haptic	138.0 (0.652)	
Audio+Video	Haptic	76.5 (0.028)	*
Audio+Video	Video	184.5 (1.000)	
Audio+Video	Video+Haptic	230.5 (1.000)	
Audio+Video+Haptic	Haptic	50.5 (0.002)	*
Audio+Video+Haptic	Video	176.0 (1.000)	
Audio+Video+Haptic	Video+Haptic	145.5 (1.000)	
Haptic	Video	77.0 (0.017)	*
Haptic	Video+Haptic	72.0 (0.020)	*
Video	Video+Haptic	241.0 (1.000)	

Table B.9: Post-hoc Wilcoxon Results — Error Count (Memory Game)

Condition A	**Condition B**	**W(p)**	**Significant**
Audio	Audio+Haptic	142.5 (1.000)	
Audio	Audio+Video	91.5 (1.000)	
Audio	Audio+Video+Haptic	81.0 (1.000)	
Audio	Haptic	78.5 (0.484)	
Audio	Video	156.5 (1.000)	
Audio	Video+Haptic	128.5 (1.000)	
Audio+Haptic	Audio+Video	93.5 (0.673)	
Audio+Haptic	Audio+Video+Haptic	91.0 (0.625)	
Audio+Haptic	Haptic	142.5 (1.000)	
Audio+Haptic	Video	153.0 (1.000)	
Audio+Haptic	Video+Haptic	105.5 (1.000)	
Audio+Video	Audio+Video+Haptic	133.5 (1.000)	
Audio+Video	Haptic	72.0 (0.033)	*
Audio+Video	Video	109.5 (1.000)	
Audio+Video	Video+Haptic	106.0 (1.000)	
Audio+Video+Haptic	Haptic	61.0 (0.043)	*
Audio+Video+Haptic	Video	96.5 (1.000)	
Audio+Video+Haptic	Video+Haptic	153.0 (1.000)	
Haptic	Video	117.5 (0.354)	
Haptic	Video+Haptic	68.5 (0.024)	*
Video	Video+Haptic	130.5 (1.000)	

Table B.10: Post-hoc Wilcoxon Results — Task Completion Time (Memory Game)

Condition A	**Condition B**	**W(p)**	**Significant**
Audio	Audio+Haptic	240.0 (1.000)	
Audio	Audio+Video	97.0 (0.026)	*
Audio	Audio+Video+Haptic	36.0 (< .001)	*
Audio	Haptic	242.0 (1.000)	
Audio	Video	58.0 (0.001)	*
Audio	Video+Haptic	86.0 (0.011)	*
Audio+Haptic	Audio+Video	118.0 (0.113)	
Audio+Haptic	Audio+Video+Haptic	86.0 (0.011)	*
Audio+Haptic	Haptic	231.0 (1.000)	
Audio+Haptic	Video	86.0 (0.011)	*
Audio+Haptic	Video+Haptic	81.0 (0.007)	*
Audio+Video	Audio+Video+Haptic	209.0 (1.000)	
Audio+Video	Haptic	158.0 (1.000)	
Audio+Video	Video	194.0 (1.000)	
Audio+Video	Video+Haptic	249.0 (1.000)	
Audio+Video+Haptic	Haptic	139.0 (0.387)	
Audio+Video+Haptic	Video	262.0 (1.000)	
Audio+Video+Haptic	Video+Haptic	227.0 (1.000)	
Haptic	Video	135.0 (0.311)	
Haptic	Video+Haptic	145.0 (0.531)	
Video	Video+Haptic	251.0 (1.000)	