DM

A Dynamic Difficulty Adjustment Framework for Serious Games Applications in Neurorehabilitation

MASTER DISSERTATION

Pedro Henrique Gonçalves Lobo
MASTER IN INFORMATICS ENGINEERING

UNIVERSIDADE da MADEIRA
A Nossa Universidade
www.uma.pt

Fevereiro | 2024

A Dynamic Difficulty Adjustment Framework for Serious Games Applications in Neurorehabilitation

MASTER DISSERTATION

Pedro Henrique Gonçalves Lobo
MASTER IN INFORMATICS ENGINEERING

ORIENTATION
Sergi Bermúdez i Badia

CO-ORIENTATION
Diogo António de Andrade Branco

UNIVERSIDADE da MADEIRA
A Nossa Universidade
www.uma.pt

Faculdade de Ciências Exatas e da Engenharia

Mestrado em Engenharia Informática

A Dynamic Difficulty Adjustment Framework for Serious Games Applications in Neurorehabilitation

Pedro Henrique Gonçalves Lobo

Orientado por:

Sergi Bermúdez i Badia

Diogo Branco

Constituição do júri de provas públicas:

Nome completo (categoria), Presidente

Nome completo (categoria), Vogal

Nome completo (categoria), Vogal

Tuesday 14th May, 2024

Resumo

Jogos sérios são frequentemente utilizados para abordar a natureza repetitiva e frustrante da reabilitação. No entanto, podem ainda cair nas mesmas armadilhas da reabilitação tradicional se o jogo não tiver desafios ou se tornar demasiado difícil. Para evitar esses resultados problemáticos, o ajuste dinâmico de dificuldade (DDA) pode ser utilizado na reabilitação não apenas para definir as configurações clínicas ideais para o treino, mas também para criar um equilíbrio entre o tédio e a ansiedade, conhecido como zona de fluxo.

Os sistemas DDA tradicionais utilizam diversas variáveis intrínsecas ao jogo e à experiência do jogador (ou seja, número de participantes, pontuação, toques na tela, tempo total para completar a tarefa) para o ajuste de dificuldade. A principal limitação desses sistemas DDA é a sua dependência de variáveis de jogo para o ajuste de dificuldade, resultando na falta de generalização para outros jogos e padronização.

Este documento explicará a primeira etapa do desenvolvimento de uma estrutura para ajudar a implementar um sistema de ajuste de dificuldade dinâmico universal e agnóstico que utiliza informações orientadas pelo paciente (sinais eletrofisiológicos, relatórios explícitos) e dados orientados pelo jogo para ajustar a dificuldade e criar uma experiência envolvente e equilibrada para o utilizador

Nesta tese, trabalhámos no desenvolvimento de um framework DDA agnóstico, que ajudará a facilitar a implementação deste sistema util em séries de jogos mais diversas, ao mesmo tempo que fornece ferramentas para ajudar os menos experientes em tecnologia a utilizá-lo.

Durante esta dissertação, foi realizado um estudo com utilizadores onde vários voluntários experimentaram três versões do mesmo jogo utilizando diferentes estímulos (explícitos, subjetivos e implícitos) para se adaptarem à dificuldade, com o objetivo de responder (i) se o sistema é capaz de criar uma experiência agradável e envolvente e (ii) se o framework pode adaptar eficazmente à dificuldade com base em vários estímulos.

Dos testes meticulosos realizados durante este projeto, surgiram algumas descobertas esclarecedoras: (i) No geral, os voluntários mostraram-se a desfrutar e a estar envolvidos durante o experimento e (ii) os voluntários mostraram-se a perceber o ajuste como orgânico e envolvente.

Este framework DDA agnóstico representa um passo significativo para otimizar os ajustes de dificuldade nos jogos sérios, levando a uma experiência de reabilitação mais eficaz e enriquecedora.

Keywords: Ajuste Dinâmico de Dificuldade · Eletromiografia · EMG · Jogos Sérios · emteqPro

Abstract

Serious games are often used to address rehabilitation's repetitive and frustrating nature. However, they can still fall into the same pitfalls as traditional rehabilitation if the game lacks challenge or becomes too difficult. To avoid these problematic outcomes, dynamic difficulty adjustment (DDA) can be used in rehabilitation not only to set the optimal clinical settings for training but also to create a balance between boredom and anxiety, known as the flow zone.

Traditional DDA systems use various variables intrinsic to the game and player experience (i.e., number of participants, score, screen touches, the total time for completing the task) for difficulty adjustment. The major limitation of these DDA systems is their dependence on game variables for difficulty adjustment, resulting in a lack of generalization to other games and standardization.

In this thesis, the Flow Optimal Framerwork (FOF), an agnostic DDA framework was developed, this framework facilitates the implementation of this system in a more diverse series of games, while maintaining an easy integration.

An user study was carried out to validate FOF where the various volunteers experimented three versions of the same game utilizing different stimuli(explicit, subjective and implicit) to adapt to the difficulty was done to answer (i) if the system is able to create an enjoyable and engaging experience and (ii) if the framework can effectively adapt to difficulty based on various stimuli.

From the testing done during this project, some insightful findings arise:(i) Volunteers overall were shown to enjoy and be engaged during the experiment, and (ii) volunteers were shown to perceive the adjustment as organic and engaging.

The Flow Optimal Framework represents a significant step towards optimizing difficulty adjustments within serious games, leading to a more effective and enriching rehabilitation experience.

Keywords: Dynamic Difficulty Adjustment · Heart Rate · Serious Games · emteqPRO · Adaptive · Biofeedback · Closed-Loop

Acknowledgements

Here now, I proudly present those to whom I owe much of my success during the making of this thesis. These remarkable people are responsible for my unrelenting willpower, discipline, and motivation. My parents and friends, and their unconditional support, my advisors and their exceptional mentorship that guided me since the inception of this thesis, their hard work made this project the best it can be. I cannot forget my colleagues with a special acknowledgment to Rodrigo Lima, whose wholehearted assistance was invaluable. They were a crucial part of collaborating on this remarkable piece of work. I sincerely thank NeuroRehabLab for their contribution in providing their facility to put my project to the test. Working in their Headquarters was not short of a gift to me, making me grow as a professional and as a human. I cannot state how thankful I am to have worked and been in the company of such a fantastic array of marvelous people. For you all, I express my gratitude.

Table of Contents

List of Figures

List of Tables

Lista de Acrónimos

AGAIN Arousal video Game AnnotatIoN

DDA Dynamic Difficulty Adjustment

EDA Electrodermal Activity

EEG Electroencephalogram

EMG Electromyograph

FET Field effect transistor

FOF Flow Optimal Framework

GEQ Game Experience Questionnaire

HR HeartRate

IMU Inertial Measurement Unit

KD Derivative Gain

KI Integral Gain

KP Proportional Gain

LED Light emitting diode

LSL Lab streaming layer

NPC Non-Playable Characters

PID Proportional Integral Derivative

PPG Photoplethysmograph

PQ Presence Questionnaire

PTM Personalize Training Module

PTP Precision Time Protocol

RESP Respiration

RQ1 Research Question1

RQ2 Research Question2

SD Standard Deviation

SKT Skin Temperature

SOPI Sense of Presence Inventory

SUS System Usability Scale

UI User Interface

VR Virtual Reality

VR Virtual Reality

1 Introduction

1.1 Context

One of the significant challenges in rehabilitation is the repetitive and often frustrating nature of exercises, which can decrease motivation and hinder progress [2]. To address this, researchers have explored the use of serious games to create enjoyable exercises that support rehabilitation [2]. These serious games integrate rehabilitation tasks with gamification elements, such as scores, leaderboards, medals and difficulty. These gamification elements aim to increase engagement and make the repetitive rehabilitation tasks more appealing. One key feature of gamification for motivation and/or frustration is a correct difficulty level. Ensuring a well-balanced difficulty adjustment is significant for gamers but arguably more crucial for patients utilizing technology. A nontailored experience might lead to the opposite effect and be suboptimal to the patient's health and experience. One way to address this issue would be from a computational perspective by introducing a Dynamic Difficulty Adjustment (DDA) system. This approach dynamically adapts the game environment and elements in real-time, tailoring the experience to the player's performance. It ensures that the game becomes more accessible or complex as needed, offering a personalized and optimal gaming experience [3].

DDA has received significant attention recently due to its potential to enhance player enjoyment and reduce negative emotions such as boredom and anxiety [4]. However, while the player's performance is often used as a metric to determine the level of challenge in the game, it does not necessarily reflect the player's enjoyment or satisfaction, as players may derive pleasure from different levels of challenge [5].

To solve this problem and increase the accuracy of the difficulty assessment, several projects added the player's emotional response as an additional metric to measure the level of challenge. In fact, some databases such as the Arousal video Game AnnotatIoN (AGAIN) that focus on the relationship between gameplay and emotional state, such as level of arousal, contains 1,100 in-game videos of multigenres games labeled for arousal by 124 participants, which goes in line with an exploration of an affect modeling tailored for participant experience [6]. Emotional response can not only be measured through subjective measures such as questionnaires but also through physiological measures such as skin temperature (SKT), electrodermal activity (EDA) or Heart Rate (HR) [7, 8]. Electrophysiology has a wide range of uses in Dynamic Difficulty Adjustment (DDA) systems. For instance, Electrodermal Activity (EDA) sensors can evaluate fear and anxiety, informing adaptive environmental changes based on the obtained results [9]. Additionally, real-time Electroencephalogram (EEG) data analysis allows for dynamic difficulty adjustments in response to the user's excitement level, as demonstrated in gaming scenarios [10]. These applications showcase the potential of electrophysiological and affect integration in DDA systems, enhancing user experiences through personalized and adaptive interactions. Finally, DDA suffers from replication issues, these systems are usually developed for a specific game leading to a new development cycle each time a new game wants to make use of this system [11].

The main goal of this dissertation is to face these limitations by creating a novel game-agnostic DDA system (Flow Optimal Framework) that can be controlled with multiple stimuli. The project structure will follow the State of the Art (Section 2) in which the scientific evidence of techniques of DDA and psychological connections will be explored. The Flow Optimal Framework (FOF) (Section 3) presentation, describing thoroughly its components and a technical validation of the system is presented. In Section 4, a case study is developed in which participants play a Unity game developed as a demo to test FOF while their physiological responses are collected and processed. Finally, in sections 5 and 6, results and discussion are presented respectively, linking the literature, limitations and future works of this work. Attachments include a manual of dependencies to install and how to implement it in other games.

In this dissertation, we introduced the Flow Optimizer Framework, a novel Dynamic Difficult System. Unlike common DDA systems, It focuses on using affective data that has been validated technically and by users. This approach helps tackle the challenges of adjusting game difficulty and addresses some of the issues found in traditional DDA systems. Furthermore, our framework will focus on providing a more user-friendly approach for the implementation, relying on real-time scripting and real-time data visualization to provide tools for the less technologically savvy. This not only enhances the overall usability of the project but also extends the reach to a larger audience.

In this dissertation, we aimed to address two principal research questions:

- **RQ1:**Is the system able to create an enjoyable and engaging experience for the participants?
- **RQ2:**Is the system able to adjust the difficulty based in different stimuli (Explicit, Subjective and Implicit)?

2 State Of The Art

In this section, the connections between Dynamic Difficulty Systems, neurorehabilitation and physiological response (biofeedback) will be explored. Exploring these topics will aid the assessment of the needs of our project, namely neurorehabilitation, to understand the needs of the framework will have to satisfy to be used in neurorehabilitation, DDA to understand how to maximize and avoid fault in the adjustment of difficulty, and biofeedback to understand which equipment we should use in this project to acquire the physiological data that will be crucial in proving the viability of the project.

2.1 Neurorehabilitation

In instances of significant disability, neurological injury, or specific conditions, neurorehabilitation emerges as a therapeutic approach. Its primary objective is to expedite recovery or, at the very least, minimize the impact of it on the person's daily life. An example of a situation that might call for neurorehabilitation would be injuries or illnesses affecting the nervous system, often resulting in impairments such as losing arm dexterity. To address these challenges, the use of neurorehabilitation in conjunction with the collaboration of a team of medical experts to guide the patient may help to improve the possibility of regaining the dexterity once lost and improve the quality of life.

One of the challenges faced by neurorehabilitation lies in the fact that a lesion to the nervous system may lead to unpredictive effects depending on where in the body the lesion occurs. It might affect multiple functions, leading to exhaustive rehabilitation programs to treat multiple issues simultaneously [12]. The brain possesses the remarkable ability to adapt to its environments. While this adaptation offers the people easiness in the process of recuperation or improvement from injuries or treatments, the process to achieve the goals requires a lot of will and time [12].

Neurorehabilitation applications can be divided into traditional and technological approaches. The traditional one is based on non-technological such as physiotherapy and paper and pencil tasks [13]. As for technological approaches, virtual reality systems have been extensively explored, using screens, projectors and head-mounted displays to provide immersive environments where the patients practice improving their motor skills, balance, and coordination [14]. Some robotic systems are also used within this context in which rehabilitation is assisted by a robot [15]. These technologies have been combined with gamification elements in which, for example, the individuals train a skill related to the injury or illness being treated while receiving score, bonus, and visual and audio feedback [16]. Using serious games with these gamification elements to help user engagement has been thoroughly explored in the literature. These systems range from physical activities, for instance, the PEPE platform containing mixed reality components in which the user controls the game actions using a Microsoft Kinect Sensor, resulting in a platform for physical activity that has been validated with aged population [17]. A neurorehabilitation cognitive tool such as Reh@city in which a virtual reality simulation of a city is presented where participants can enroll in various activities similar to daily living experiences [18]. A further study that the usage of Reh@City led to higher improvements in cognitive domains when compared with tasks based on paper and pencil [13]. Finally, programs such as NeuroAIreh@b add gamification and task programs led by artificial intelligence systems that help tailor the rehabilitation programs to participants' needs [19].

2.1.1 Key features for Motivation

In the context of ensuring competent neurorehabilitation, it is crucial to consider key features that contribute to both motor learning and motivation, drawing parallels with traditional video games' emphasis on engagement. These features focus on two topics: motor learning, the learning process of better performing a specific motor skill through training and motivation [20]. The objective of the motor learning key feature is to maximize improvement in each session. On the other hand, motivation key features focus on maintaining the patient's interest and engagement through rehabilitation so it does not hinder progress (Fig. 1).

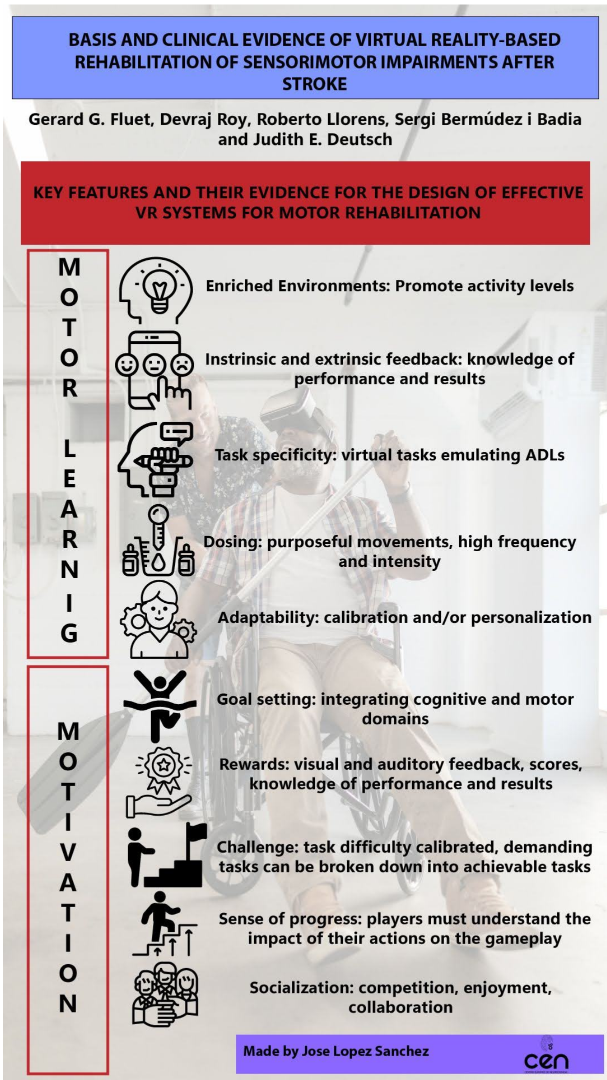


Fig. 1: Motor Rehabilitation Key Features [20].

- **Motivation**

  - **Goal Settings:** In Virtual Reality Rehabilitation, clear and achievable goals are key to leading the patients to an engaging experience. Breaking tasks into smaller, more manageable goals will also lead to a higher sense of accomplishment by the patients. A goal focuses on attention in a specific movement, can reinforce the neural pathways, and intensifies concentration. Goals can also serve as a measure of the progress of the patients, which can be used as a guide to improve the rehabilitation process [21].

– **Reward:**In VR Rehabilitation the reward, has important role to play, to motivate the patients to go beyond what normally would. These rewards could range from visual and auditory feedback to knowledge from your performance, like scores and results. Rewards are important to avoid boredom and fatigue in an environment where these feelings are natural. [22].

– **Challenge:** In this context, it refers to the creation of a task that is meant to make the patient surpass their current limits. When faced with a challenging task, the brain is stimulated to create a new neural connection. This process called neuroplasticity [23] is essential for the creation of motor recovery [24].

– **Sense of Progress:** Crucial element in VR rehabilitation for stroke patients. These elements can be defined to give the patient an understanding of their impact action. which is important to avoid the lack of motivation due to the feeling of stagnation. There are many ways to create a sense of progress like scores, difficulty adaptability [23].

- **Socialization:** The incorporation of social interaction in a VR environment as a way to motivate and improve the patient's mood and progress in rehabilitation. These features can take many forms, like collaborations and competitions [25].

Maintaining player engagement presents a challenge, as extremes of frustration or easiness can hinder the gaming experience; a balanced approach is ideal. This challenge can be tackled from a computational point of view by implementing a Dynamic Difficulty Adjustment (DDA) system. DDA makes use of key gameplay variables to create a gaming experience both balanced and exciting to the player.

2.2 Dynamic Difficulty Adjustment

DDA is a system that can be implemented to create a gaming experience that adapts to the player, to create a constant and appropriate challenge to the player, intending to create a more exciting and balanced experience. This section of the work will delve into the research of DDA: the primary problems that affect DDA, how to assess player skills correctly, how to implement changes in difficulty, and approaches for measurement of player skills and implementing the adjustment of difficulty [4].

2.2.1 DDA in Rehabilitation

One typical use of the DDA is to increase the engagement of the patients during rehabilitation. Rehabilitation is an arduous and repetitive process that halts improvement, and the patient, if the patient is not concentrated or motivated, has a solution for this project. So to avoid this undesirable outcome, DDA is used to remedy it because it provides a way to create a fair challenge that stimulates the player's engagement and motivation. Examples of DDA system in rehabilitation:

- **Rehabilitation Gaming System:** A virtual reality-based neurorehabilitation tool designed for stroke recovery by utilizing the brain's plasticity. The key component of this system is the Spheroids, which are rehabilitation focus scenarios where the patients are tasked to capture or place a sphere that moves toward the patient (Fig. 2) [26]. In this project dynamic difficulty was implemented using a Personalize training module which is created based on the effect of the training parameter on the player during each scenario.The training of this module was made by two groups of people, one of 21 acute/subacute stroke patients and one control group of 20 healthy people.

Fig. 2: Rehabilitation Gaming System [27].

The Personalize Training module can adjust the difficulty for each trial independently and automatically. To adjust the difficulty, this module uses the following steps (Fig. 3). Firstly, even before the beginning of the training, a calibration task is to be done to define a baseline that will be used throughout the project. Secondly, after each of the ten trials, the PTM will modify the difficulty level based on the performance. The gaming system will create an accompanying gaming parameter for each new difficulty level recorded. It takes into account the user's prior responses and the psychometric model of Spheroids. The difficulty level increases by 10% after each time the patient intercepts 70% or more spheres during the trials. The difficulty will rise until it reaches the maximum difficulty of 100%. On the other hand, the difficulty will decrease by 5% each time half or fewer spheres are intercepted. Hence, the game parameter constantly changes to match the patient's performance.

- **Fruit Catcher:** An exercise game where the player has to perform a weight shift lateral exercise to catch, with a basket, the apples that fall from a tree (Fig. 4). This game uses the modification of the range (the scope of positions that the apple can fall from), speed (how quickly the apple falls from the tree), and accuracy(how big the basket is) based on the player's last session raw data. In conjunction with raw data used to calculate the player's skill, an AI is implemented that will ask the player for his emotional and physical well-being to adjust the game based on it to avoid the over-extensions of the player [28].

- **Happy Fish:** A serious game that uses a robotic device has a bridge between the patients and the game and assists the player with the movements. In this game, you control a fish that needs to dodge the upcoming wall of seaweed (Fig. 5). The difficulty is adjusted based on the player's ability to achieve the target score and alters the change that the fish need to cover and the velocity of the fish [29].

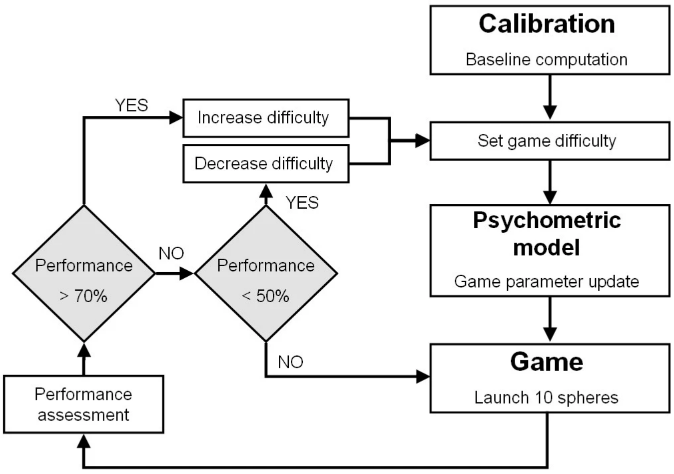


Fig. 3: Flow diagram of the RGS Personalized Training Module [26].

Fig. 4: Fruit Catcher Serious Game [28].

DDA-MAPEKit framework:A solution for Unity3D using DDA this framework allows the usage of multiple DDA algorithms. It was tested in an exergame called CicloExergame (Fig. 6), in which players had to pedal on a cycle ergometer to move the player, while real-time biofeedback regarding their heart rate and blood oxygen saturation was provided on the screen [30]. This framework, developed for Unity Engine, focuses on easing implementing different DDA strategies in the same game, by enabling customized treatment for each game mechanic. The customization for each mechanic is done by creating a different MAPE-K loop tied to each one of them. This project although a proof concept was shown to achieve positive results.

VR Exergames: Created with the objective of boosting exercise motivation where the players has to traverse a maze (Fig. 7) procedural generated which includes a series of exercise to complete. This project utilizes DDA as a way to maintain motivation of player over time by creating more fine tuned experience [1].In this project the DDA system was paired up with experience driven Procedural Content Generation, to create a optimal level for the skill that the individual as at the moment. To create the ideal difficulty at the end of each maze the player are asked to rate how difficult and exhausting it was in a 5-point Likert scale (1-not at all, and 5-extremely) .

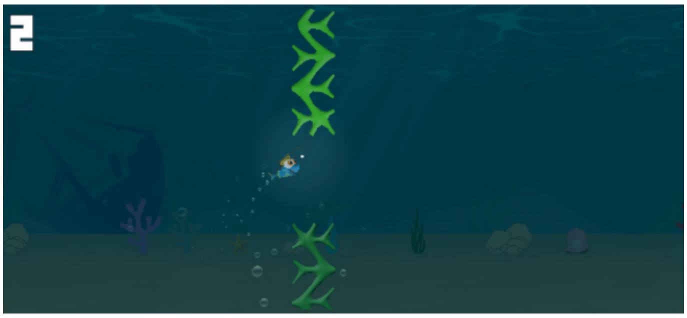


Fig. 5: Happy Fish Serious Game [29].

Fig. 6: DDA-MAPEKit Game [30].

2.2.2 Difficulty Adjustment Approaches

As vital for DDA as measuring the player skills, if not more, the calculation of the adjustment is one of the most significant and more complex steps in implementing a DDA system. In this section of the project, we will explain some of the most common approaches to calculating the adjustment needed for the game [3, 4]:

- **Reinforcement Learning:** This technique is used by training an AI based on positive reinforcement, in which the AI is rewarded when taking action toward a desired outcome. In this case, the reward would be the in- increase in player performance, and the action would be the adaptation of the game towards a balanced difficulty. This technique requires a lot of programming, data analysis, and interaction to collect enough training data to be effective [31].
- **Probabilistic methods:** This technique uses probability theory to model and analyze the player's performance and to predict the best changes to create balanced gameplay. It requires much complexity and training data to perform at its full potential [4].
- **Single and multi-layered Perceptrons:** Uses Artificial Neural Networks to make decisions based on the data collected from past players to adjust the difficulty. Due to their faster and more straightforward nature, single-layered Perceptrons are used in a more linear game with few features. In contrast, multi-layered Perceptrons are complex more powerful, and better equipped to deal with more complex games [32].
- **Upper Confidence Bound for Trees and Artificial Neural Networks:** Uses the Decision tree or Artificial Neural Networks algorithm to predict based on collected data. These algorithms learn from past player data to make decisions to predict with a confidence inter-

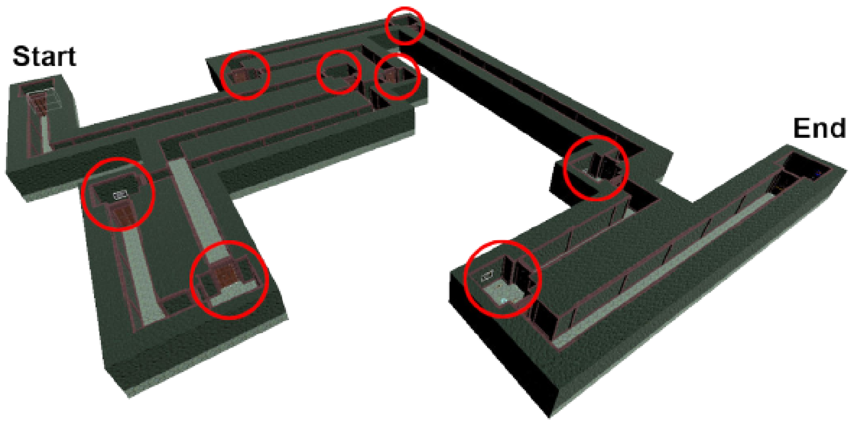


Fig. 7: VR Exergames

val to remedy the unknown factors. These techniques are complex to implement and require significant data analysis to work at total capacity [33].

- **Hamlet System:** Uses a Neural Network to create a model of the player's skills based on the past data and the metrics defined beforehand to predict the future states of the player. When an unwanted state is expected, the system will tweak the game settings to prevent it. This technique will require an elevated amount of data and might not adapt in real time but will over time [3].
- **PID Controller:** The Proportional Integral Derivative (PID) controller is a complex system of feedback control that has vast uses in different applications, from engineering to industrial applications, to help regulate the values of variables through feedback. This system manipulates the system or personal feedback by strategically adjusting specific aspects based on system or personal feedback. It seeks to influence and optimize the system or individual by modifying relevant parameters to achieve a specific setpoint.
- **Dynamic Scripting:** Uses online machine learning to alter the behavior of game elements. The behavior of each game element will change based on rules defined beforehand, each rule as an associated weight which will determine the change of activating the rule [34].
- **Self-organizing System using Artificial Neural Networks:** This artificial neural network uses unsupervised, which means the data is not labeled, and the algorithm is not told what is wrong or correct and is up to the algorithm to find connections between its data by itself to determine the optimum difficulty for the player. This technique is complex to implement and requires significant data analysis to work at total capacity [35].
- **Biofeedback:** Biofeedback consists of techniques that use sensors to record real-time data about a person's physiology, like brain waves, muscle tension, heart rate, or skin temperature, and provide real-time feedback that can be used to improve the person's health. These techniques can help many treatments, like assisting people in learning relaxation techniques [36,37]. Depending on the health condition and the goal of the treatment, the kind of biofeedback might change to be more accurate or less physically taxing. The more commonly used variety of sensors used in biofeedback include the following:

– **Electromyography(EMG):** Monitors the muscle activity, commonly used to diagnosticate muscle injury and nerve damage [38, 39].

– **Skin Temperature(SKT):** Monitors the skin temperature and can be used in training to reduce headaches in conjunction with the EMG [40].

– **Electroencephalography (EEG)**: Monitors the scalp for brain activity, commonly used to diagnose epilepsy [41], or to measure levels of attention [42, 43] or engagement [44].

– **Photoplethysmography (PPG):** Monitors the blood flow in the microvasculature of skin, commonly used in smartwatches to monitor the heartbeat [45].

– **Respiration(RESP):** Monitors the abdominal and chest movement while breathing, usually used in respiratory treatment and anesthesia [39].

2.2.3 Challenges: Implementation

Getting a good measurement of the player skills challenges means the success or failure of the DDA system. According to [8], DDA approaches to measure the difficulty can be divided into three categories:

– **Formal model:** A model that entails a mathematical description of the game rules to help predict and analyze the players' next move. Using a game tree enables a quick transition between states to reach one that matches the player's skills [8].

  - **Intervention:** Turn completion.
  - **Advantages:** The same algorithm can be used in different games.
  - **Disadvantages:** Needs a detailed model of the game.
  - **Recommended:** Board games and logic games.

– **Selected Features:** Used when the formal model is unavailable or too complex to implement. This approach uses two types of game features as the base of its prediction. These are parameters that describe selected game elements (nº enemies, nº lost lives) or Indicators that tell us about the player's performance (kills per second, coins per second) [8].

  - **Intervention:** Depend on the game's structure.
  - **Advantages:** Capability of using a set of features to implement DDA on a complex game
  - **Disadvantages:** One game can have multiple different parameters and performance indicators.
  - **Recommended:** Arcade games, shooter games, racing games, platformers.

– **Direct Examination:** Measures the difficulty through the feeling the player perceives during the gameplay. Data about the player's feelings can be extracted by querying the player or measuring the player's emotions during the game. This approach can be used in conjunction with the other two [8].

  - **Intervention:** Depend on the game's structure.
  - **Advantages:** Has access to a direct source of how the player feels.

- **Disadvantages:** To measure the player's feelings, there is a need for a tool that might disrupt the player.

- **Recommended:** Arcade games, shooter games, racing games, platformers.

The implementation consists of mathematical functions that determine how much the values of the variables should be changed to suit the player's skill. Data should be collected first so the mathematical operations can be defined and done multiple times. At the beginning of the game, the difficulty should be based on data recorded during testing before any data on the player has been gathered. One effective way to collect data early is to use a tutorial so that the player does not start with an inappropriate challenge for them [4].

During the implementation, one of the main concerns is to make the transition from the wrong difficulty to the right one as seamless as possible so as not to disrupt or overwhelm the player. To help with this, the difficulty should be updated in small, subtle increments. No two or more changes should be made simultaneously on the same object to avoid problems or the need to change the gameplay drastically. If drastic changes are necessary, they should be done off-screen, so the player does not notice them. To accurately adapt to the game's difficulty, the best approach is to create a mathematical function that uses data collected from multiple tests to calculate how much the variables need to change [4].

During assessing a player's skills, one of the critical steps is to select the performance indicators that will be used to determine the player's performance. Each variable will be linked to a game parameter, such as the score-to-time ratio connected to the game parameter number of coins. These game parameters can be divided into three sections [4]:

- **Attributes:** It involves altering specific game parameters to adjust the mold difficulty to the desired shape. Although it is the more accessible and straightforward way to enhance the difficulty, altering the values of attributes of the game has the disadvantage of being too unrestrained, leading to one variable having a vast number of attributes, making it hard to choose which one is fitted. Information is critical to alleviating this troublesome issue, depending on how the collection data of game parameters is triggered. If it is collected after an event, for example, "Level completion", we can use the information gathered to determine if the player took too much time and if it was due to nº of the enemy size of the map. On the other hand, spikes in the data can determine how to adjust the difficulty if the variables have permanent data collection [4].

- **Behaviours:** It consists of the actions that non-playable characters (NPC) and other game elements can take. This can enhance the difficulty by increasing or decreasing the complexity of the character's behavior to match the player's skill. For example, in a stealth game, the enemy can have more complex movement patterns to create difficulty for a skilled player trying not to get detected [4].

- **Event:** It consists of planned occurrences triggered by some criteria being met. Easy to implement, it comes at the cost of needing planning before implementation, so it is not too noticeable to the player. One example of an event will be an enemy dropping a health item if the player has low life [4].

To effectively implement DDA in a game, it is essential to carefully consider which approach will be most suitable for the specific type of game. This will help to ensure that the difficulty level is accurately calculated and that the implementation does not introduce unnecessary complexity.

2.2.4 Challenges: Assessing Player Skills

The first step the DDA system needs to consider is to evaluate the player's skill, so the game can fully know how much it must change to reduce anxiety or boredom. This is done by comparing the current performance with the expected performance [4].

- **Variable:** Also known as performance indicators, these are the values used to evaluate the player's performance. These values must clearly represent a desired or undesired outcome and if the player's skills are improving [4].
- **Data Collection:** Describes when the variables should be updated. Data collection can be done event-triggered, where you need a specific event to occur to update the variables, or permanent is done in a game loop where from x to x time, the variable will be updated [4].
- **Reference Point:** This point of reference will be compared to the values collected during the data collection to determine if the difficulty should change. Using personal beliefs to create the reference point is a suboptimal way to do it since it would rely on guessing the correct value. Another option is to program an AI agent to play the game at a medium level of difficulty multiple times to collect worthy and reliable data. However, the best way to collect an appropriate reference point is to use real player playthrough data that had an enjoyable experience, due to the data collected from this reference being more realistic and needing fewer attempts [4].
- **Data Analysis:** It is the part where we compare the Variables and Reference points for each one to determine if the level of challenge is optimal and what should be done to achieve the ideal balance. Evaluating the player should be done in true and false or skills ratio. Data Analysis should be done constantly so the game difficulty can change at the right time and make it more difficult for the player to notice the change, so it does not disrupt the player [4].

Implementing DDA in a game is a multifaceted task that relies on the careful consideration of a diverse series of factors. Carefully choosing the approach, determining which values to use as a stand-in for the skills of volunteers in the implementation, and deciding how to evaluate these skills are crucial for the most faithful adjustment possible. Even when treated with the attention it deserves, DDA systems may still be plagued by inherent problems of this type of system.

2.2.5 Problems

Although DDA undoubtedly offers advantages, as mentioned before, it has key drawbacks that discourage its use. This section of the document will explain the limitations of implementing and using a DDA system. The problems that may be faced are the following:

- **Difficulty balancing:** The definition of the incorrect variables to evaluate the player's performance may lead to a joyless and unbalanced experience [46].
- **Player feedback:** Players might have different perspectives on what makes an appropriate challenge. So it is necessary to receive feedback about the DDA system from many types of players [47].

- **Complexity:** This type of system requires complexity due to the continuous analysis and updating needed to maintain optimal difficulty [48].
- **Testing and debugging:** The system needs to be tested thoroughly to ensure that the difficulty is always up to par with the player's challenge [48].
- **Integration with other game systems:** It needs to be integrated with another game system to work at its maximum capabilities, like game AI or level designer [48].
- **Game-Specific Tailoring:** DDA systems are tailored to specific games, making it challenging to implement them in other games [11].
- **Player Frustration:** If the difficulty is adapted too quickly too many times, it may disrupt the player due to feelings of unfairness or boredom [46].
- **Lack of Control:** The player's experience might get disrupted if the player lacks control due to changes in difficulty without being prepared [4].
- **Inconsistency:** If the changes are too quick or too many, it might result in an inconsistent and joyless experience [4].
- **Linearity:** Some players might prefer a linear difficulty progression instead of being constantly challenged [48].

The implementation of DDA presents numerous challenges, from inherent issues to its system, including the necessity for thorough testing and the requirement for tailoring the system to specific games, to problems associated with a flawed implementation of the system, including the frustration of the players or the disruption of the experience. Despite these barriers potentially hindering the system's adoption, the ability to tailor the gaming experience for each player justifies the struggle required.

DDA systems are useful systems to break the monotony inherent to rehabilitation, creating a better experience for the patient, which will improve the progress done while in it. But although its a rather valuable system, its implementation comes with a series of steps to certify the most accurate results, from choosing how and which information to converting it to difficulty. One implementation plan for this project is the direct examination, which will include biofeedback to examine the feelings of the player to better acquire a more compatible difficulty.

2.3 Conclusion

To conclude this section, we can confirm that in this exploration, through the scientific fields relevant to this thesis, we found valuable insights that will be used to create a solid foundation. Comprehending the specific needs of neurorehabilitation, optimizing difficulty adjustments, and picking suitable biofeedback equipment are key components that will allow us to achieve the best result possible.

3 Framework Proposal

This section introduces the Flow Optimizer Framework (FOF), a game-agnostic framework. FOF leverages physiological signals to autonomously adapt to game difficulty. The framework employs a hybrid approach, utilizing Unity3D and PyFlow (`https://github.com/wonderworks-software/PyFlow`) for data visualization and processing, along with Lab Streaming Layer (`https://github.com/sccn/labstreaminglayer`) for network connections. This section will explore the framework's structure, technical validation, and a case study demonstrating the practical application of FOF.

3.1 Structure

The Flow Optimizer Framework was made to address one of the biggest problems with DDA, which is the restrictive nature of the system, as it is usually only compatible with one individual game. This restriction hinders its use, as it is a one-off solution and may not be worth the costly effort of implementing. FOF was created as a simple and open-minded approach to DDA to solve this.

FOF will take on the challenge of facilitating the use of the framework for non-technical users by using a real-time visual scripting tool and visual tool. As it will help ease the use of the framework as it permits people who do not understand programming, in general, to interact with the DDA system without the risk of unintentionally breaking the code. The real-time feature of the system will be a valuable help as the user can receive immediate feedback on their changes and adjust the values.

Another feature implemented to facilitate the project was the ability to display the data received in the application of a dynamic graph known as PyMonitor to help give the user a clearer view of the transitions of data during use, making it easier to detect any abnormal behavior with the system.

This framework also seeks to help in neurorehabilitation, where serious games are a mainstay in the fight against boredom and sameness that commonly pains the rehabilitation. Nevertheless, a static difficulty may restringe the game's effectiveness even with serious games. To avoid this, FOFs will focus on fulfilling some key features (Fig.1) of neurorehabilitation, these being:

- **Goal Setting:** The player can set a simple goal by using the difficulty as a metric of progress.
- **Rewards:** Sense of accomplishment after increasing the difficulty to a higher degree.
- **Challenge:** The DDA system will help create an adequate challenge for the player, making it feel challenged without frustration.
- **Sense of progress:** The adjustment of difficulty will progress with the progress of the player's performance, so the player can measure his progress by looking at how difficult the game has become.
- **Socialization:** Patients who played the game may enjoy comparing how high the game's difficulty got with each other.

During the execution of this Framework, a PID Controller (DDA algorithm) will be paired up with multiple stimuli to showcase the ability of this system to adapt to various scenarios. The different stimuli used during this project were restricted to three:

– **Explicit:** Stimuli refer to a game's obvious and easily identifiable elements that present cues to the player intended to prompt a particular response. Example: the score of a game, lives of the player.

– **Subjective:** Refers to cues that are perceived differently depending on each individual's personal thoughts and feelings. One such example is a questionnaire inquiring about the stress or enjoyment of the player during the game.

– **Implicit:** Refers to cues that influence the player without them realizing it. Example: sound cues that play when the player makes a mistake.

The Flow Optimizer Framework has many components, the most relevant being Unity framework, Python application, Physiological Unit and the Lab Streaming Layer framework, as shown in the diagram below:(Fig. 8).

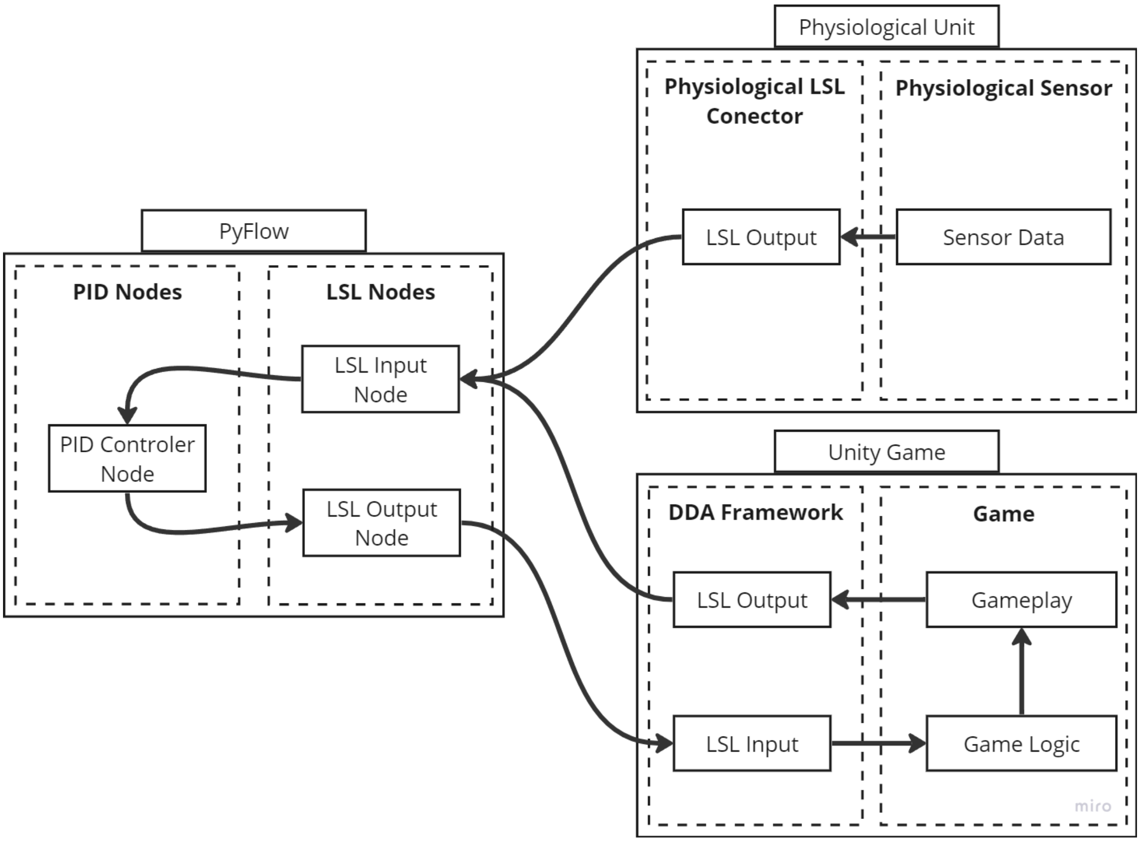


Fig. 8: Proposed Framework for the Project.

3.1.1 Lab Streaming Layer

Unifying the many parts of this Flow Optimizer framework, the Lab streaming layer (LSL) framework is an open-source framework created to facilitate the gathering of almost real-time measurement data during research systems. This tool is precious due to the ability to manage networks, time synchronization, real-time access, and optionally centralize data collection, viewing, and disk recording. The LSL system was ideal for connecting the different systems that made up this project.

Stream discovery, data transmission, and time synchronization protocols are at the core of the LSL framework [49]. The stream discovery protocol is used to discover all streams available without any significant explicit configuration. Used to send the streams the frameworks rely on to spread, this framework uses the data transmission protocol, which incorporates extensive descriptive metadata and a simple encoding format. The time-synchronization protocol is used to accurately ensure the real-time alignment of the data sent from different sources. The synchronizing of each stream is done by calculating the offset by utilizing a subset of the Precision Time Protocol (PTP) algorithm to calculate clock offsets.

An advantage of this framework is its ability to be used in a significant range of languages compatible with the system, including the main languages used to create the Flow Optimizer framework, such as C# and Python [50]. The flexibility of this framework is all due to the LSL library. The library comprises the core transport library called liblsl and the language interfaces (Fig. 9).

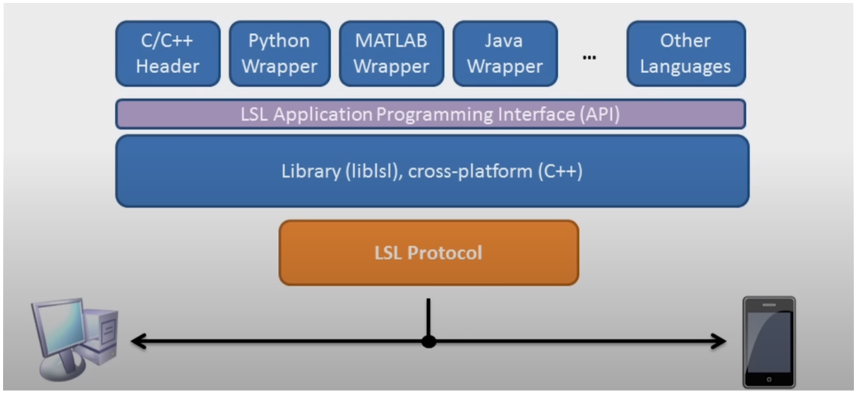


Fig. 9: LSL Library

3.1.2 Unity Apps

Unity's unique position in the game development industry and its extensive ecosystem of tools and resources make it an exceptional choice for our project. We access many pre-built components, plugins, and community-generated assets by anchoring our framework on this rich platform. This resource increase will significantly reduce development time and streamline the integration of fundamental functionalities. Moreover, Unity's general adoption ensures compatibility with various projects, making its use of all the systems more generic.

The Unity framework (Fig. 11) was developed as a generic way of easing data transmission and reception through LSL. An integral feature of the system entails the ability to designate any script and select variables for data transmission and storage (Fig. 10). Furthermore, the framework should allow the user to wield control over the data stream's starting, pausing, and ending and allow the user to change key stream characteristics, for example, name, sample rate, and channel count.

3.1.3 Physiological Unit

In response to the project's demand for incorporating physiological data from a real-time sensor and the framework exclusively designed for LabStreamingLayer, a new component of this project will be

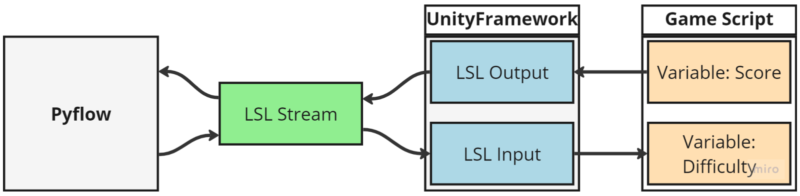


Fig. 10: Unity Diagram

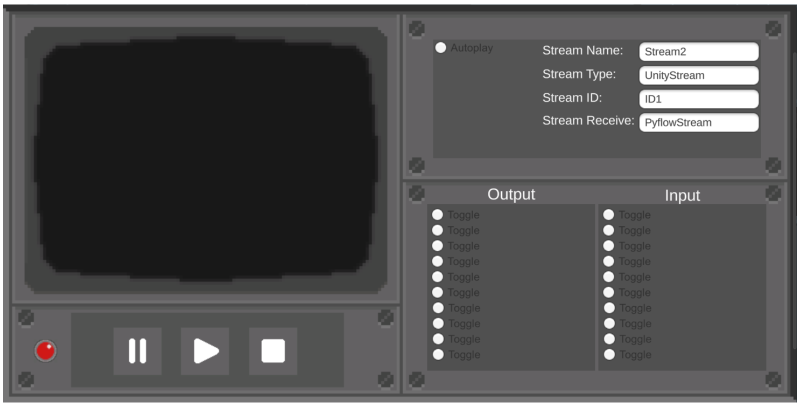


Fig. 11: DDA Framework

needed to certify the transmission of the data through LSL to the framework. This "Physiological Unit" component will bridge the sensor data collected and the DDA Framework. The data collected by the sensor has a crucial role in showcasing the project's capability to utilize a myriad series of real-time stimuli to guide the adjustment of the difficulty of the games.

3.1.4 Python App

While the sole reliance on a Unity approach for the framework might have performed well, it would have some constraints and presented significant limitations in its applicability for neurorehabilitation settings.

Python was selected because of its flexibility and the availability of various libraries, such as Pyqt5, for effective data processing. It offers finer control over data compared to Unity. Additionally, Python allows us to use the visual scripting tool Pyflow (`https://github.com/wonderworks-software/PyFlow`), making it user-friendly for non-IT professionals to adjust as needed. To ensure the FOF's success, it is essential to establish a seamless flow of data between the Unity project and Python. Data will be sent from Unity to Python and adapted using an algorithm to dispatch back to the Unity game with a reasonable amount of delay. Furthermore, a specific visualizer/monitor, PyMonitor, was developed to visualize the received data in real time, handling multi-threading and multiple LSL streams.

This complex system will be divided into four significant sections (Fig. 12):

- **PyMonitor**: This section displays all the real-time data received in a series of graphs.
- **Data Receiver (Single Stream Grapher)**: Through LSL, this section of the python will need to be able to receive data.

- **DDA Algorithm (PID Controller)**: This section will focus on creating the ideal difficulty for the player based on the data received.
- **Data Dispatcher(Stream Transmitter)**: Through LSL, this section will output the new difficulty calculated.

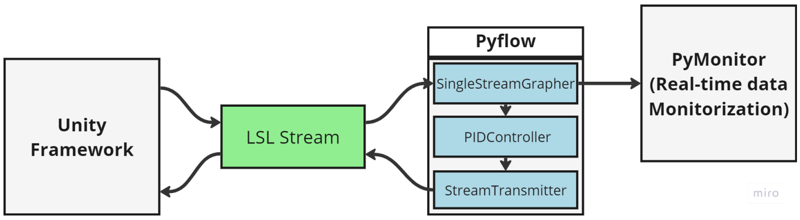


Fig. 12: Python Diagram

To facilitate the use of the FOF for individuals with limited or no experience in game engines and to streamline the process of testing various parameters within the DDA algorithm, Pyflow will be employed as the solution. Pyflow represents a Python-based visual scripting tool that enables the creation of intricate programs by utilizing boxes, referred to as nodes, each containing succinct sections of code that will use pins to send and receive data from other nodes. When these nodes are combined, they allow us to generate intricate programs that possess a visually coherent structure, consequently enhancing user-friendliness, particularly for those with limited expertise (Fig. 13). Another valuable feature of this framework is the ability to change values while a program is running without stopping or restarting the program to make it easier to test a variety of values for each parameter of the people and the ability to create new custom nodes and pins to make the program more fit to what the user needs.

Some new nodes had to be created to make the Pyflow fit better to calculate DDA. These nodes can be divided into four major sections: LSL, DDA, Sensors, and Helpers.

LSL nodes: These have the function of controlling and managing the LSL streams created to send data between the major components of the DDA system. The nodes are the following (Fig. 14):

- **SingleStreamReceiver:** Receives data from a single stream and sends the last one second of data received to the nodes connected to it.
- **SingleStreamGrapher:** Receives data from a single stream and sends the last one second of data received to the nodes connected. A new process will also show the data graphically using the PyMonitor.
- **SingleStreamSample:** Receives data from a single stream with a single channel and sends it the value when its value differs from the previous value.
- **MultiStreamReceiver:** Receives data from all streams available and sends the last one second of data received from each stream to the nodes connected to it.
- **MultiStreamGrapher:** Receives data from all streams available and sends the last one second of data received from each stream to the nodes connected to it. A new process will also show the data graphically using the PyMonitor.

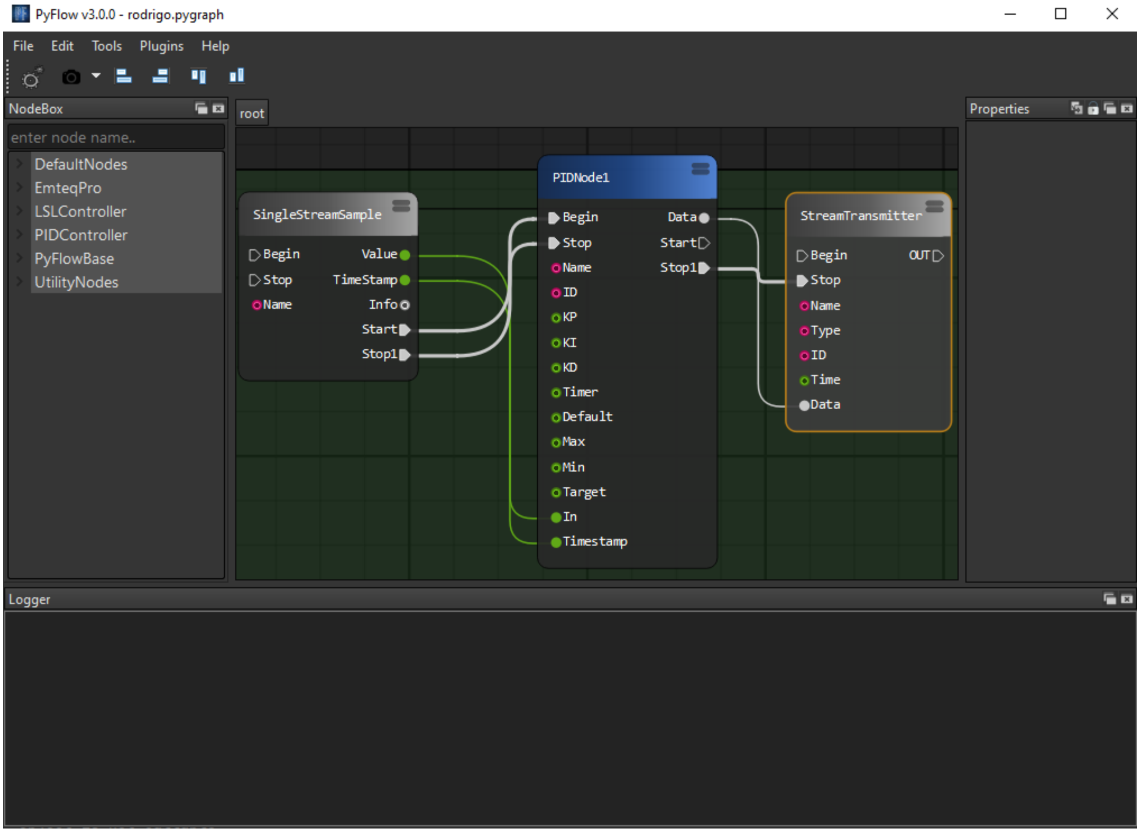


Fig. 13: Pyflow Example.

– **StreamTransmitter:** Establishes an LSL data stream designed to transmit the information the node receives.

– **StreamGrapher:** Establishes an LSL data stream designed to transmit the information the node receives. It will also show the data sent in graphs using the PyMonitor.

– **PIDController:** This Node is used the PID algorithm (Fig. 15a) to calculate, without prior data during the player playthrough, the optimal changes in difficulty for the player based on a SetPoint is the objective that the player should be, the performance of the player and the constant KP, KI and KD that help the arrival to the correct difficulty and to impede overshooting. This controller can be used with all types of variables, making it easy to use and generic, but with the caveat of testing much value until the user finds the correct constant for the user-specific situation. It also creates a JSON file in which the data it uses for help with analyses of the data.

Utility Nodes: This node was created to support creating accessible and optimized programs that use the mentioned nodes. The nodes are the following (Fig. 15b):

– **SaveJSON:** Save data received by the PID Controller in the JSON file.

– **Data Dispenser:** Node with the function of filtering the data of Dictionary and lets the user retrieve all the data as an array, or just the last value received, giving more considerable flexibility in using this node.

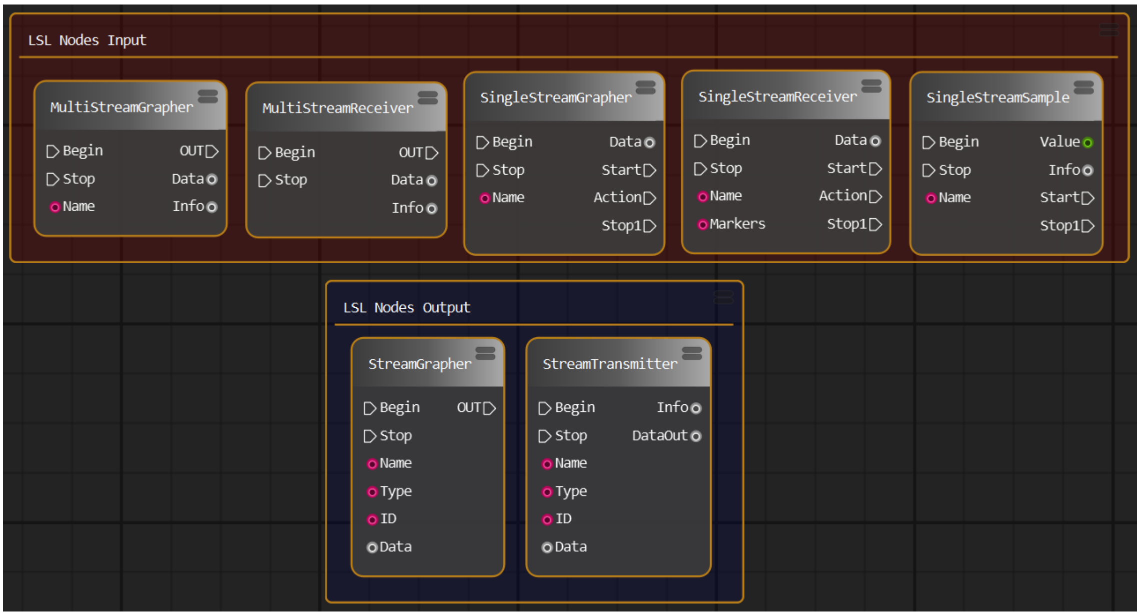
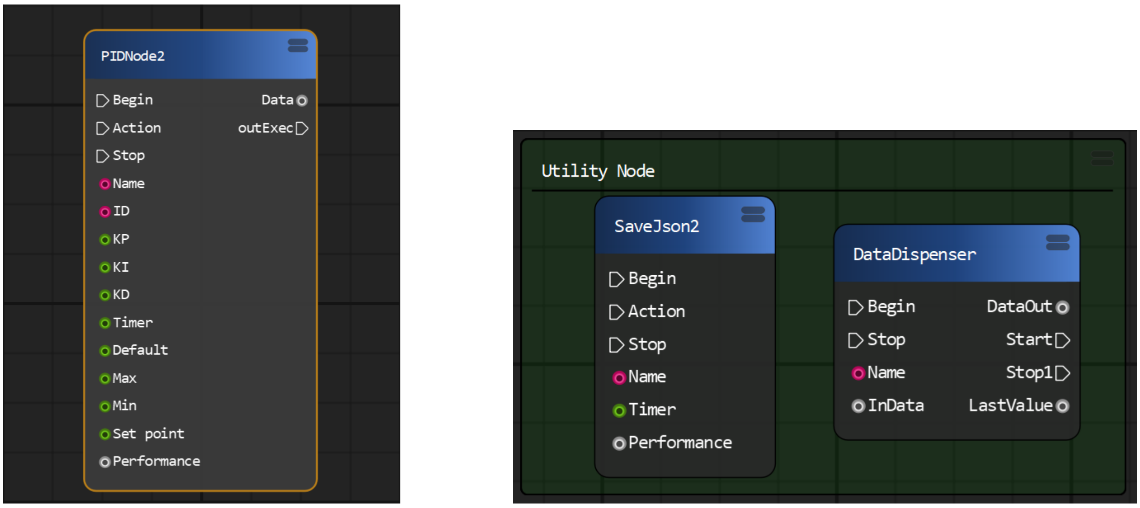


Fig. 14: LSL Nodes.



(a) PID Node

(b) Utility Nodes

Fig. 15: Utility & PID Nodes.

3.1.5 PyMonitor

This application was created using the Python libraries PyQt5 and PyQtGraph, enabling the creation of graphs using data received from an LSL stream. The primary objective is to provide a clearer picture of the data being worked on (Fig. 16). The application's workflow kickstarts with the initialization of the user interface (UI) and the graphs configured based on the information from the LSL stream. Once everything is set up, the program runs in an endless cycle. During this loop, it continuously updates the graphs as it receives data from the LSL stream, adhering to its frequency rate. One of the most significant challenges faced during the implementation of this application was ensuring a consistent flow of data to prevent the graphs from displaying stuttering or jagged lines while using multi-threading to avoid the PyMonitor performance affecting the speed of Pyflow. This issue needed to be solved without compromising accuracy and integrity. The solution involved creating a buffer that stores one second of data, which is gradually reduced while maintaining a smooth and uninterrupted visual representation and using memorization, which consists of storing and reusing the results of functions to lower the frequently used functions. Additional features were implemented to ease the uses of the PyMonitor, which are the following:

– Ability to save and load UI settings, to facilitate faster graph setup.

– Ability to filter the graph by sources, to assist users in locating the specific data they need.

– Ability to choose the source of each channel, to provide users with enhanced control over the content displayed by the PyMonitor.

– Ability to add or delete graphs(max 10), to display only the relevant data.

– Ability to change the color and width of the graph's line, to enhance the clarity and comprehensibility of the graphs.

Theses chosen extra features were added based on other similar programs like OpenSignals [?], StreamViewer [?] and BrainVision LSL Viewer [?].

3.2 Technical Validation

Software, like everything, has limits, and knowing them is good practice to avoid unfortunate events like crashes or data loss. In a research project like this one [1], those events quickly erode the project's data fidelity, destroying the credibility of the project overall. To avoid these undesirable outcomes, one of the most critical steps of this project is to test the limits of the system involved with the project, namely the Pyflow, PyMonitor and Unity LSL framework.

3.2.1 Methodology

To mitigate the possible undesirables, we will test the project thoroughly by following a series of meticulous and focus group tests that will take the system to its extreme, giving us the complete picture of the project limits. The tests are the following:

– **Performance Stress Test:** Checked if the system could handle the load without significant delays or crashes. It was tested by generating a large amount of data from Unity and sending it to Python. Measured the time it took for the data to be transferred, processed, and returned to Unity.

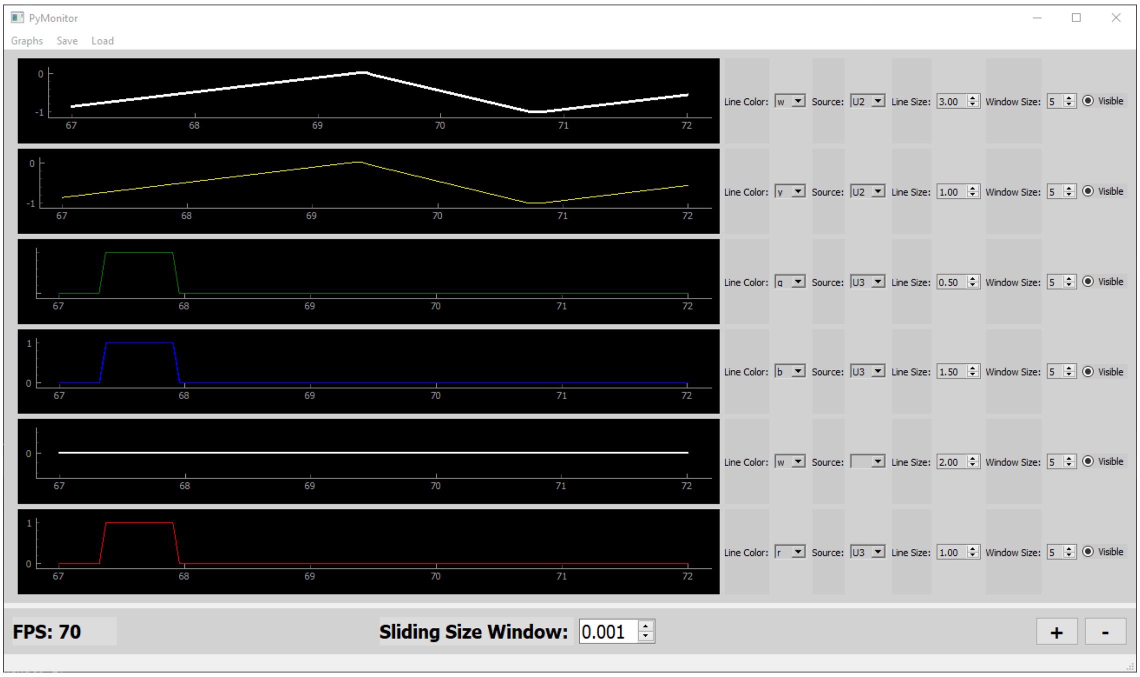


Fig. 16: PyMonitor.

– **Real-time Data Test:** Evaluated the system's performance when dealing with real-time data streams. Tested by sending continuous data from Unity to Python (e.g., sensor data, player movements), processing it, visualizing it in the graph, and ensuring the graph updated smoothly and accurately.

– **Boundary Testing:** Tested the system with extreme values and edge cases. It was tested by sending very small or large data sets, negative values, or data exceeding the expected range. Ensured that the system could handle such scenarios gracefully.

– **Long Duration Test:** Monitored for memory leaks or performance degradation over time. It was tested by running the system for an extended period, continuously sending data back and forth between Unity and Python.

– **Multi-User Test:** Assessed if the system could handle multiple connections without data corruption or interference. It was tested by simulating multiple users, sending data simultaneously from separate Unity instances to the Python script.

3.2.2 Evaluation Criteria

Even with the best tests possible, the test would only be relevant with suitable metrics to quantify the results. For these tests, we created metrics that test each system individually and as a complete system to clarify the project limits to its zenith. The metrics we decided to use were the following:

– **Execution Time** We will measure the program's performance at different time intervals (10s, 60s, 300s, 600s, 1200s) to assess its behavior over extended periods.

– **Success Rate** This metric will determine whether the program remains operational after a specific duration (X time).

- **Synchronization** We will measure the time taken for data sent from Unity to be displayed in the PyMonitor. This will help us evaluate the efficiency of data transmission.
- **Consistency** We will assess the occurrence of stuttering in the PyMonitor and categorize it as follows:
  - **Very Frequent (0)** Indicates stuttering happens frequently.
  - **Common (1)** Indicates stuttering occurs occasionally.
  - **Rarely or None (2)** Indicates stuttering is infrequent or absent.
- **Data Reception Delay** This metric will measure the time it takes to receive data sent from Unity, providing insights into data transfer efficiency.

3.2.3 Setup

Before starting the experiment, a default setup should be discussed to maximize the velocity and efficacy of the study. For this project, 3 PCs are required, which shall be used to run all the software that is a must in this project. The software used for this project will be the following:

- **Testing Program:** Unity program explicitly made to send and receive the stream that will be sent that will be used to evaluate the Execution Time, Success Rate, Consistency and Data Reception Delay (Fig. 17).
- **Stream generator**: A simple unity program created specifically for this project to send a customizable LSL stream to help push the system to its limit.
- **Pyflow:** This program, in conjunction with the Testing program, is the pillar of this test; this system will receive, display, and send back the streams sent by the previously mentioned software.
- **OpenSignals [51]:**This program is used in specific case scenarios of the experiment to see how the system handles real-time data. The sensor used with this program was the biosignalsplux.

Fig. 17: Test Program.

The PCs utilize during this technical validation will have the following specs:

- **First PC:**This desktop was equipped with Windows 10, an Intel i7-6700 3.40 GHz processor with 16 GB of RAM.
- **Second PC:**This desktop was equipped with Windows 10, an AMD Ryzen 5 3550H 2.10 GHz processor with 8 GB of RAM and an AMD Radeon Vega 8 graphics card.
- **Third PC:**This desktop was equipped with Windows 11, an Ryzen 5 3600, GTX 1070 ti, and 16GB RAM.

In its entirety, the system will be meticulously configured using the detailed schematic provided. This schematic is a blueprint for the system's configuration, providing a clear and well-structured configuration (Fig.18).

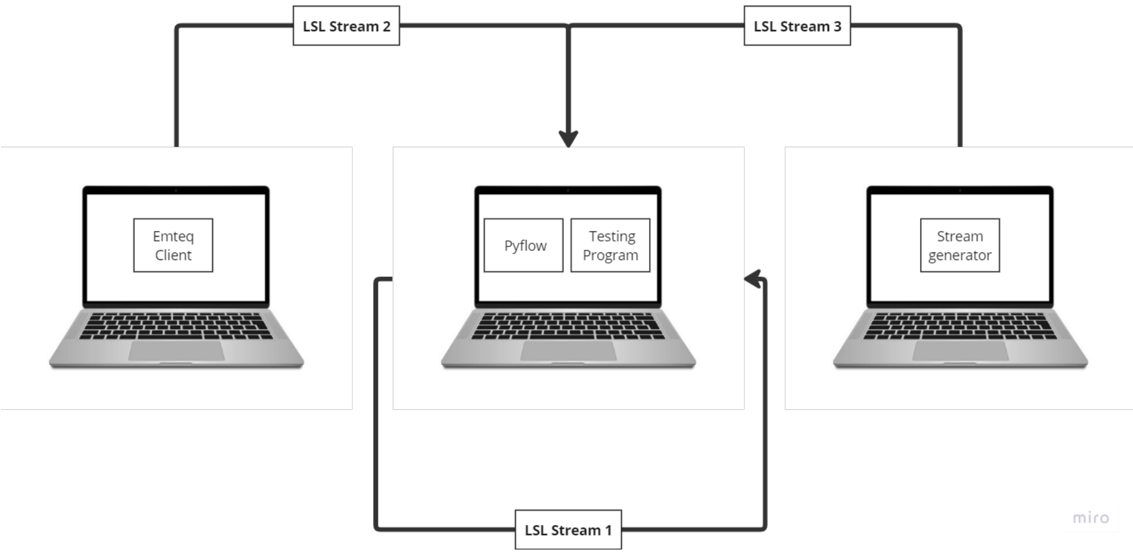


Fig. 18: Software Test Setup.

3.2.4 Test Cases

To ensure that the entire project was tested thoroughly with the methodology, several settings were chosen to more easily identify the cause of any error, bug, or crash that might happen during the several experiments. The setting is Streams, which includes the number of LSL Streams running while testing each case scenario, channels combining a number of all data channels sent in every stream; Sample Rate is the number of values sent every second by a specific stream and PCs which are the number of computers that will be utilized in each case of the experiment. The scenarios that will be followed will help identify the limits the system will be following (Table.1)

3.2.5 Results

This experiment focuses on ascertaining the limits, weaknesses, and strengths of the Pyflow or the PyMonitor to determine its overall utility. The objective is to identify the optimal and most effective ways to utilize these tools.

For each evaluation criterion established in the system assessment, a series of statistical metrics, including averages and standard deviations, were calculated to achieve a nuanced interpretation of

Table 1: Test Cases.

Nº	Streams	Channels	Sample Rate 1	Sample Rate 2	Sample Rate 3	PCs	Sensors
1	1	1	100	0	0	2	0
2	1	7	5000	0	0	2	0
3	2	17	100	100	0	2	0
4	2	17	5000	100	0	2	0
5	2	27	5000	5000	5000	2	0
6	3	3	100	100	100	2	1
7	3	10	1000	1000	1000	2	1
8	3	27	5000	100	100	2	0
9	3	19	5000	4000	100	3	1
10	3	19	5000	5000	5000	3	0

the results. The Success Rate resulted in straightforward outcomes, with an average value of 1 and no deviations. Meaning that each scenario was seen through until the end without any crashes.

In contrast, synchronization had an average of 4.9 seconds and a substantial standard deviation of 10.9. To address the major variation in values, a recalculation was conducted without the two outliers, resulting in an average of 1.4 seconds and a standard deviation of 0.03. This change was made to accommodate the presence of the higher outlier and consolidate the rest of the values.

Moving on to Consistency, the metric had an average value of 1.74 and 0.56 deviation. The deviation of this metric only appeared due to the decline of the performance, exclusive to the last two experiments.

The Data Reception Delay had an average value of 0.11 and a standard deviation of 0.04. Suggesting a relatively stable performance in data reception, with a minimal deviation throughout all scenarios.

The result of our study indicates that the system's capabilities are up to the challenge imposed by this dissertation's needs. This conclusion was reached due to the results in the table(Table 2) and graphs(Fig.19, Fig.20, Fig.21) below, which demonstrate that the system can operate with high data values clearly in the fifth scenario, handle real-time data from the sensors as seen in the seventh, sixth scenarios and maintain the consistent visual clarity of data in the PyMonitor across most scenarios.

Table 2: Technical Validation Results

Nº	Successful	Sync	Consistency	Delay
1	1	1.51	2	0.07
2	1	1.45	2	0.12
3	1	1.43	2	0.11
4	1	1.44	2	0.21
5	1	1.39	2	0.12
6	1	1.43	2	0.06
7	1	1.32	2	0.10
8	1	1.43	2	0.09
9	1	35.86	0.4	0.10
10	1	1.41	1	0.13

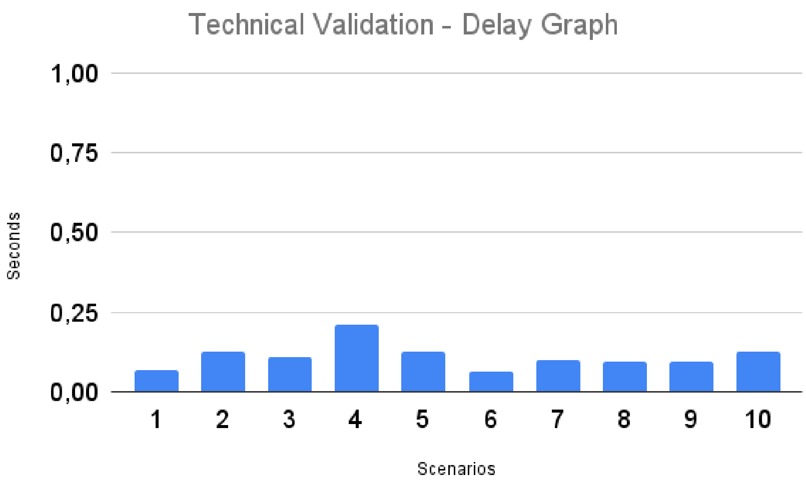


Fig. 19: Technical Validation - Delay Graph

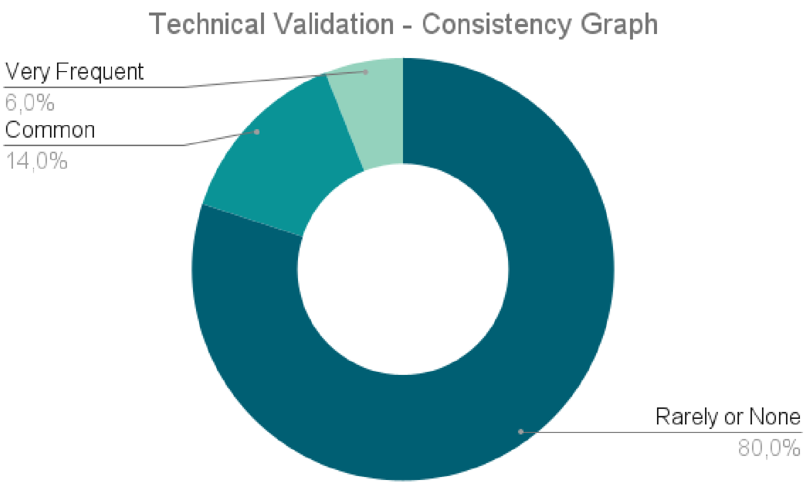


Fig. 20: Technical Validation - Consistency Graph

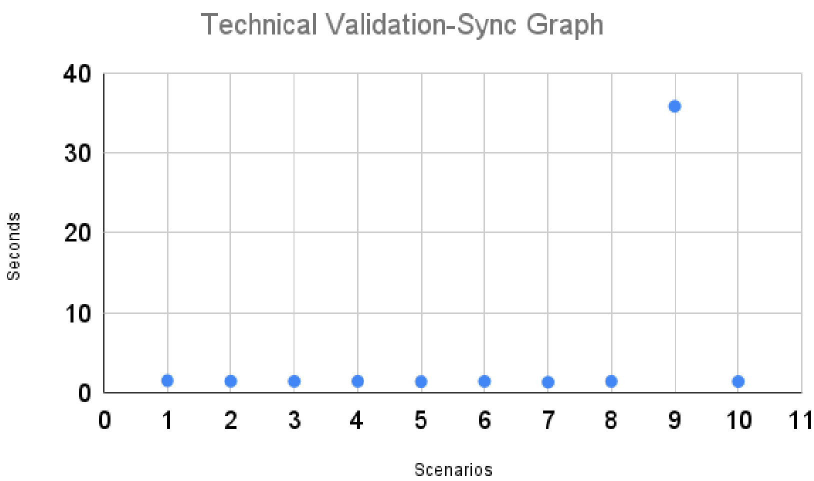


Fig. 21: Technical Validation-Synchronization Graph

However, with some case scenarios of the project, namely the ninth and tenth scenarios, it can determine an apparent loss in performance from the PyMonitor, visual stuttering being more prominent, and the delay increase in showing the changes in data graphically as can be seen clearly in the ninth scenario especially, theses faults may be due to the use of boundary values in which the system is not able due to handling the mass number of data sent from different PCs in the case of the tenth scenario, and due to the significant difference between the sample rates in each stream used in the ninth. Nevertheless, the Pyflow was not displayed to be compromised as the PyMonitor was in these scenarios.

Another thing to note about this trial is the loss of data clarity in the PyMonitor after the minimization of this program, which will persist for moments and return to normal after some seconds.

3.3 Conclusion

FOF is a unique system that aims to rectify prevalent issues within DDA systems by creating an agnostic framework that will supply a simple and efficient way to adjust most games' difficulty dynamically while also supplying tools to help the less knowledgeable use the project. The framework was made of various components; the DDA unity framework and the physiological unit will focus on retrieving data from the game and receiving the adjusted difficulty, while the Python apps will use a real-time open-source visual scripting tool named Pyflow to run a DDA program created to received the data from the games and alter the difficulty based on it. Following the detailed description of project concept components, a rigorous test was conducted to test the system's efficacy, which ended with the conclusion that the system was fit for the needs of the dissertation.

4 Case Study

Within this dissertation segment, we will discuss the methodology by which the project will be subject to verify its worth and success. We will start by explaining all the systems created to validate the project. Subsequently, an in-depth analysis of how the algorithm was perfected to create the most engaging difficulty. Finally, a comprehensive plan for executing the user study will be outlined, serving as the conclusive examination to verify the project's positive results.

4.1 Validation

The validation of FOF will be centered around two programs created to serve as examples of the use and capabilities. One of these programs will be a VR game created with Unity3D to showcase the FOF's ability to adjust game difficulty automatically based on different stimuli.

The second program will be an interface between the FOF and the essential biofeedback necessary to test the FOF's ability to adjust the difficulty based on real-time physiological data. For this purpose, the emteqPRO face mask system [52]was chosen as the source of biofeedback. This system offers a diverse set of physiological data, including EMG for reading facial expressions, recording skin conduction during stimuli to detect arousal and detecting the changes of the valence when subject to stimuli through the measure of specific facial muscles located checks, forehead, and eyelids, PPG that can calculate heart rate, heart rate variability, breathing rate, and changes, and Inertial Measurement Unit(IMU) uses an accelerator and gyroscope to track the user's head movements.

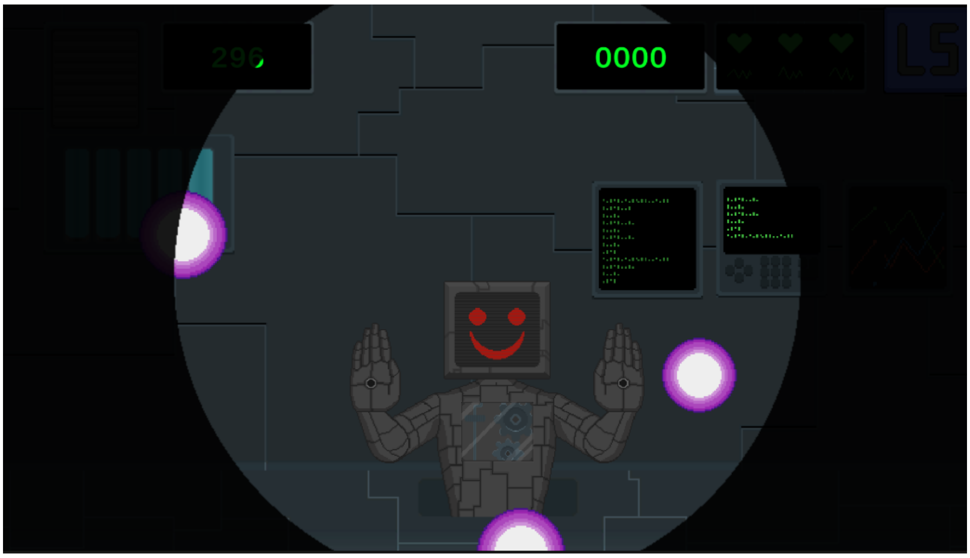
4.1.1 Unity Game

Before the framework can be tested in serious neurorehabilitation games, it must be tested in simple-to-understand games where difficulty adjustment can be tested easily and quickly. A simple game was developed titled "Robot Mania".The game was developed in Unity (Unity Technologies, 2022.1.9f1) to assess and test the capabilities of FOF and the usability of the resulting adaptive experiences created for the user. In this game, the player has to destroy balls that spawn out of the hands of a robot. The spawn location alternates between the right and left hand; thus, every second, each hand has a 50% chance of spawning a ball. The game has two versions: one designed for PC screens(Fig. 22a), where players use a mouse for interaction, and another version tailored for head mounted displays headsets(Fig. 22b). For this project, the HTC VIVE PRO EYE was used as it allows the players input to be controlled using eye-tracking in addiction to allow seamless integration of EmteqPro.

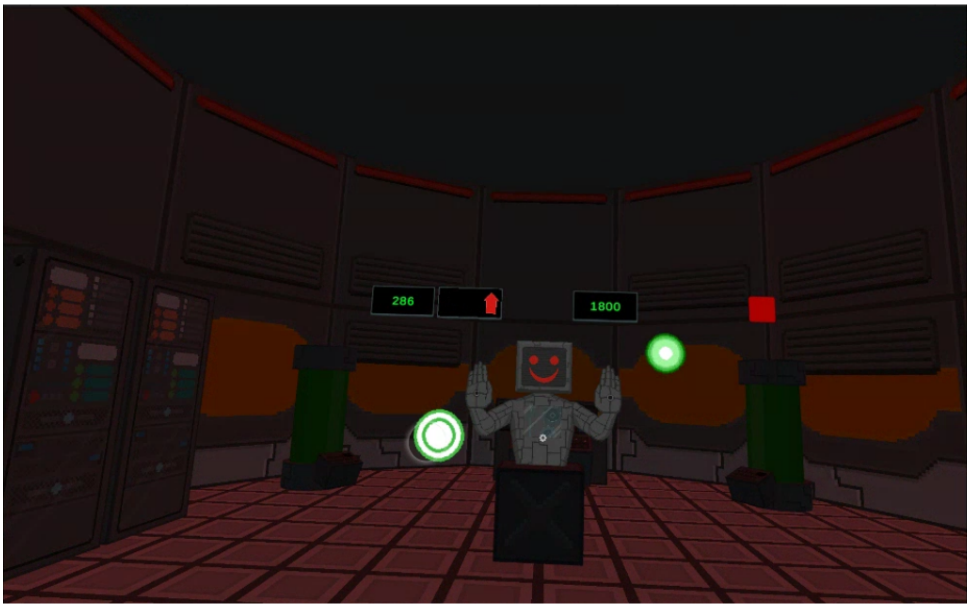
It must be implemented before the game can be used in the DDA framework. There is a need for some information about the game to be defined. Performance Inputs, Parameters, and Reference Points must be chosen so the DDA can work at maximum capacity.

Performance inputs are crucial to determine the changes needed for the game to adjust the difficulty due to needing to be more accurate. The following performance inputs are the ones chosen to classify the player's skill in the game :

– **ErrorRate:** The percentage of missed balls in the last ten balls. It will be used to determine the performance of the player explicitly. The Reference point that will be used will be 0.5.



(a) Desktop Version



(b) VR Version

Fig. 22: RobotMania

- **HeartRate:** The rhythm of the heartbeats of the player. It will be used to determine the performance of the player's simplicity. The Reference point that will be used will be the baseline of the player plus 20
- **Questionnaire:** Question from one to five where the player tells us how concentrated he is. It will be used to determine the performance of the player explicitly. The Reference point that will be used will be the 3.

After choosing the variables used to calculate the player's performance, he will need to choose how the difficulty is going to adapt to the different values the variable can take:

- **Field of view:** How much the screen shows during the game. For example, depending on the player's heart rate, the Field of view will increase and decrease. This value was be updated every 200ms.
- **BallsSpeed:** The speed of each game ball. For example, depending on the ErrorRate or Questionnaire, the speed of the ball will increase or decrease. This value was updated every 1s.

4.1.2 Emteq Client

Used as the Physiological Unit of the system, the Emteq Client(Fig. 23), will be used to send the data from the myriad of sensors that make up the emteqPRO face mask using the LSL. To facilitate the data gathering and to avoid sending unnecessary data, the Client will let the user choose which specific data the user can send. The several signals that can be sent through the application are the following: Heart Rate Data: Data retrieved with information about the current, average and deviation of the user's heartbeat. The Emteq Vr Manager SDK calculates this information.

- **EMG Data:** Data retrieved from the seven facial EMG that exist in the emteqPRO. The precise data retrieved are left and right Zygomaticus, Orbicularis, Frontalis and center.
- **PPG Data:** Data retrieved from the PPG sensor found at the center of the mask. The data retrieved will be the compound of the PPG signal and its proximity.
- **Heart Rate Data:** Data of the current, average and deviation of the heart rate of the wearer of the mask. These physiological signals are retrieved using SDK, which uses the PPG data to calculate them.

4.2 Framework Parameter Optimization

The PID Controller is a crucial pillar in the dynamic difficulty calculation of this project. As a result, it is essential to determine the most optimal constants to adjust the difficulty, ultimately enhancing the player's experience. Multiple experiments were conducted on the PID Controller to achieve this objective using different values for its constants. It is important to note that while the resulting values may be optimal for one person, this system will only adapt perfectly to everyone with proper testing beforehand due to the nature of the PID controller.

4.2.1 Methodology

A comprehensive testing procedure was executed to reach the best attune to the maximum degree of the PID controller for each stimulus, which is explicit, implicit, and subjective. Each constant was thoroughly evaluated in multiple trials to ensure that the best constants were selected for optimal performance.

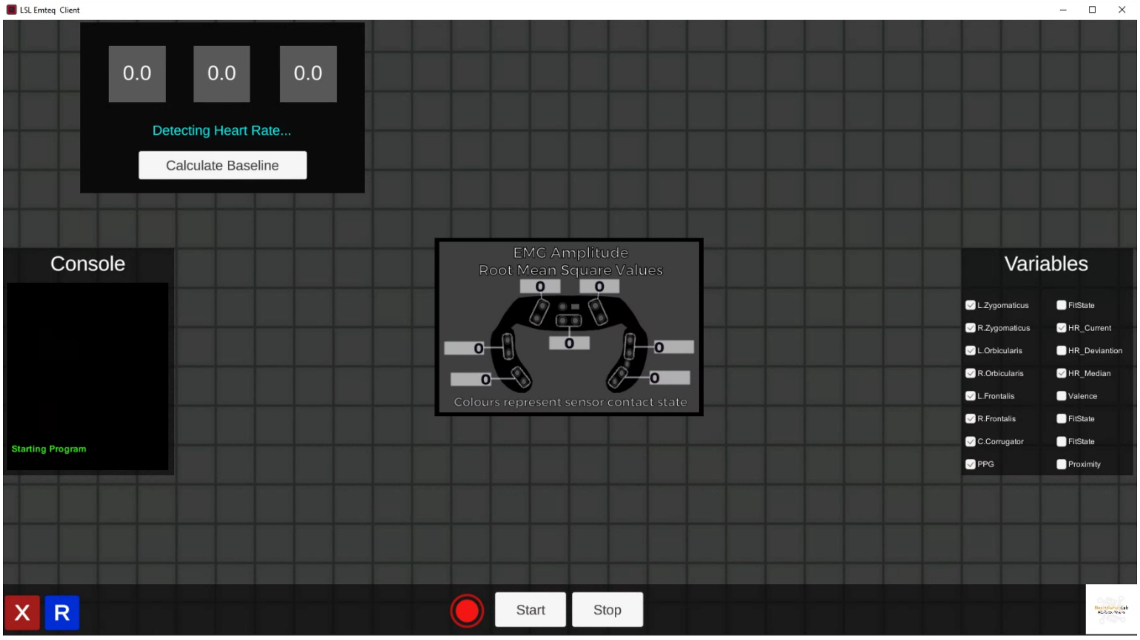


Fig. 23: Emteq LSL Client.

To evaluate the optimal constant for the PID, we will use the average of the absolute error of that result after each trial. Based on the average error, another value for the constant will be tested.

In each trial, we will test a constant value in multiple combinations multiple times to gather a robust dataset, rewarding us with a comprehensive understanding of each parameter's performance and choosing the most adequate value.

For each stimulus, we will choose three significant variables: performance, which is the variable that shows how the player is doing; setpoint of the ideal performance that the player should have; and output, which is the variable that will impact the game. For each stimulus will be a performance, setpoint and output as seen in table3.

Table 3: Stimulus.

Condition	**Implicit**	**Explicit**	**Subjective**
Performance	Based on participant's heart rate (HR)	Based on player's performance	Based on self-reported stress level
SetPoint	20% higher than baseline HR	50% of balls caught	Middle point of the Stress Scale
Output	Field of View	Ball Speed	Ball Speed

4.2.2 Results

To more accurately adjust the difficulty to a point in which most volunteers acquire the closest to the PID Controller setpoint, the constants KP, KI, and KD have to be experimented with until we find the ideal values for each one for each PID Controller, we plan to use the project. To do this, we will experiment with each constant multiple times for each PID and identifying which value resulted in the lesser average absolute error (performance-setpoint). The results acquired during this test were the following:

Explicit

This area will examine the constants to reach the values that lead the volunteer to reach the setpoint defined for the explicit stimulus, a ball failure rate of 50%, as frequently as possible.

KP: The value arrived at during the experimentation to provide the most suitable results for the constant KP was 0.1, as witnessed in the following table 4.

Table 4: Explicit KP.

KP	Experiment nº1	Experiment nº2	Experiment nº3	Average
0.010	0.239	0.266	0.265	0.257
0.050	0.123	0.128	0.147	0.133
0.100	0.008	0.086	0.154	0.083
0.125	0.150	0.129	0.126	0.135
0.150	0.071	0.150	0.145	0.135
0.200	0.139	0.143	0.147	0.143

KI: The most suitable value for the constant KI was determined to be 0.05 during experimentation, as evidenced by the accompanying table.5.

Table 5: Explicit KI.

KI	Experiment nº1	Experiment nº2	Experiment nº3	Average
0.010	0.151	0.155	0.13	0.145
0.050	0.12	0.129	0.111	0.12
0.075	0.137	0.122	0.16	0.139
0.100	0.129	0.136	0.12	0.128
0.125	0.139	0.145	0.139	0.141
0.150	0.162	0.136	0.132	0.143
0.200	0.13	0.13	0.127	0.129

KD: After careful experimentation, we determined that the constant KD value of 0.05 consistently produced the most optimal results, as illustrated in the attached table.6.

Table 6: Explicit KD.

KD	Experiment nº1	Experiment nº2	Experiment nº3	Average
0.010	0.137	0.108	0.141	0.129
0.025	0.122	0.143	0.145	0.137
0.037	0.113	0.122	0.127	0.121
0.050	0.147	0.101	0.115	0.121
0.100	0.155	0.173	0.14	0.156

When we finish the experiment, we infer that the most suitable values for the constants of this setup of the PID Controller would be KP = 0.1, KI = 0.05, KD = 0.05.

Subjective

KP: Our experimental analysis concluded that the constant KP value of 0.1 consistently achieved the most favorable outcomes, as evidenced by the accompanying table.7.

Table 7: Subjective KP.

KP	Experiment nº1	Experiment nº2	Experiment nº3	Average
0.005	0.191	0.264	0.361	0.207
0.007	0.285	0.247	0.718	0.314
0.009	0.287	0.29	0.127	0.264
0.010	0.182	0.183	0.425	0.2
0.100	1.213	1.059	1.197	0.892

KI: Due to the high values found during the experimentation with this constant, it was decided to nullify the value. In other words, the KI of this PID Controller will be 0.

KD: After careful experimentation, we determined that the constant KD value of 0.003 consistently produced the most optimal results, as illustrated in the table.8 .

Table 8: Subjective KD.

KD	Experiment nº1	Experiment nº2	Experiment nº3	Average
0.001	0.222	0.451	0.203	0.292
0.002	0.349	0.218	0.453	0.340
0.003	0.192	0.21	0.365	0.256
0.004	0.194	0.187	0.445	0.275
0.005	0.215	0.418	0.201	0.278

When we finish the experiment, we infer that the correct values for the constant of this rendition of the PID Controller would be KP = 0.009, KI = 0, KD = 0.003.

Implicit

KP: The value arrived at during the experimentation to provide the most suitable results for the constant KP was 0.1, as witnessed in the following table 9.

Table 9: Implicit KP.

KP	Experiment nº1	Experiment nº2	Experiment nº3	Average
0.001	3.042	2.857	2.857	2.918
0.002	5.045	6.086	4,366	5.166
0.005	6.135	4.293	6,129	5.519
0.007	6.867	4.193	5.779	5.613
0.010	3.456	2.691	4.777	3.641

KI: Due to the consistently high values observed during experimentation with this constant, it was determined to set its value to zero. In other words, the KI of this PID controller will be 0.

KD: After careful experimentation, we determined that the constant KD value of 0.05 consistently produced the most optimal results, as illustrated in the table.10.

Table 10: Implicit KD.

KD	Experiment nº1	Experiment nº2	Experiment nº3	Average
0.001	6.418	6.825	6.371	6.538
0.005	4.341	6.142	4.467	4.983
0.010	4.769	4.062	4.400	4.410
0.015	3.869	3.846	3.091	3.602
0.020	5.830	5.426	4.273	5.176

After the experiment, we infer that the correct values for the constant of this rendition of the PID Controller would be KP = 0.009, KI = 0, KD = 0.003.

4.3 User Study

This project section will explain how the usability test will be organized to clean and quickly obtain the most relevant data possible. The participant time during the experiment will be divided between being the subject of three experiments and answering questionnaires that seek to know how the player felt about the experiments and how they make them feel.

4.3.1 Methodology

This section of the document will explore the methods used to assess the Flow Optimizer Framework ability to create engaging and fun experiences for the volunteers.

In this user study, each participant will be presented with three versions (Subjective, Implicit and Explicit) of the same game that use different stimuli that dictate the flow of difficulty in each game. These stimuli are identical to those used during the optimization of framework parameters. (table.3).

4.3.1.1 Transversal data collection

This data was collected in all experiments for further direct comparisons.

- **PID Controller data:** The data collected from the PID Controller that was used during the respective experiment.
- **In-Game data:** The data collected from the game itself will include, score, time, score per second, Reflexes, ball speed, error rate of the volunteer.
- **Extra data:** The data collected from the stimulus not used during the experiments.

The results of the experiments will be considered positive if the data shows that the Flow Optimizer framework was able to adapt to each player using the different stimuli mentioned before.

In conjunction with all versions of the game proposed, the volunteers will also be subject to various questionnaires providing insight into their feelings through the experiments. The questionnaires used in this study were the following:

Games Background Form: This form will be used to understand the player's background in the world of the game due to the possibility of it affecting the performance and how the adjustment of difficulty will interact with the player((see section 8.2)). This questionnaire was specifically designed for this project.

Game Experience Questionnaire: To better understand the player's experience with each version of the game, the Game Experience Questionnaire will be used as a validated tool to measure the player's feelings towards the game he or she just played. This questionnaire will have two versions, core and post-game, focusing on the experience during and after the game [53]. Each question in each questionnaire aims to identify critical factors that influence the player's overall gaming experience, these are (see sections 8.3 and 8.4):

- GEQ-Core:
  - **Competence:** This factor is crucial for understanding how skillful the player felt during the game, how high their mastery and achievement related to the game was, and whether the difficulty was appropriate for the player.
  - **Sensory and Imaginative Immersion:** Use to determine how the player engages in the game environment and gameplay.
  - **Flow:** Determines if a player enters a state of "Flow" where the focus is completely gathered in the game, losing the notion of time and if yes, to what degree.
  - **Tension/Annoyance:** Determines the level of tension and annoyance that the player felt during the game.
  - **Challenge:** Determines match is the player's skill and the challenge created by the game if the player felt that the challenge provided was fair and if it motivated the player to perform better.
  - **Negative affect:**Did the player feel to any extent any negative emotion during the game, and what elements contributed to the player's negative emotions?
  - **Positive affect:** Did the player feel to any extent any positive emotion during the game, and what elements contributed to the player's positive emotions?
- GEQ-Post-Game:
  - **Positive experience:** This component evaluates if the player remembers the game experience fondly and what it caused the good experience.
  - **Negative experience:** This component evaluates if the player remembers the game experience badly and what caused the bad experience.
  - **Tiredness:** This component indicates how much the game experience eroded the player's energy and the level of physical and mental effort necessary for the game.
  - **Returning to Reality:** This component will determine the lingering feelings and effects of playing the game.

DDA Questionnaire: Difficulty adjustment is a core part of this thesis. To figure out how well it was implemented in this project, some questions asked to the player to determine how natural, consistent, fast, and engaging the adjustment of difficulty was during the play through each test (see section 8.5).

Version Comparison Questionnaire: Due to the feeling of the player changing after playing the different game versions, it is necessary to ask different questions to analyze which version exceeds or performs in different aspects. Each section question will be followed by another question for the player to justify the last answer(see section 8.6).

System Usability Scale Questionnaire: Also known as SUS, this questionnaire was created to analyze the usability of a system, product, or service. To calculate the result of this questionnaire, we sum all ten questions and then multiply by 2.5, and the closer the score is to 100, the more usable the system is. Based on research on the SUS questionnaire, if the score that results from the study is below 68, the usability is below average, and if it's 69 or above, it is considered above average (see section 8.7) [54].

Sense of Presence Inventory Questionnaire: This questionnaire, also known as SOPI, is one most common tools to measure an individual subjective experience in a virtual reality environment. It uses four categories to evaluate the virtual experience, these being the following (see section 8.8) [55]:

- **Spacial Presence:** The degree to which the games disconnect the player from the real world and connect to the virtual world.
- **Engagement:** The degree to which the VR environment keeps the player emotionally and mentally engaged in the environment.
- **Ecological Validity / Naturalness:** It measures the level at which the player feels the game replicates aspects of the real world.
- **Negative Effects:**Did the VR environment cause any negative emotion during the game, and what elements contributed to the player's negative emotions?

Presence Questionnaire: Focus on evaluating the feeling of presence inside virtual environments; this questionnaire will give us a better idea of the realism of experience by evaluating the following aspects of presence inside a virtual environment (see section 8.9) [56].

- **Realism:** Refers to the degree to which the virtual environment is perceived as lifelike, convincing, and similar to the real world.
- **Possibility to act:** Assesses how individuals feel they can interact with and manipulate objects or elements within the virtual environment.
- **Quality of interface:**The Quality of interface evaluates the effectiveness and user-friendliness of the tools or devices used to interact with the virtual environment.
- **Possibility to examine:**This component assesses the extent to which users can explore and examine the details of the virtual environment.
- **Self-evaluation of performance:**This component assesses the extent to which users can explore and examine the details of the virtual environment.

The questionnaire results are successful if the player appraises the project positively in both the Game experience and dynamic difficulty adjustment questionnaire in each experiment, showing high results in every positive characteristic, such as example, competence and flow and low results in the negative characteristics, for example, tension and tiredness for most volunteers of the project. All the other questionnaires will be used to evaluate how well-designed the system was and how good the UX was.

4.3.2 Experimental Procedure

To ensure successful project outcomes, several tasks must be completed before commencement. Firstly, each participant will be assigned a unique ID to differentiate results. The second step involves conducting baseline heart rate measurements using the Polar H10 Heart Rate Sensor (Fig. 25f). These measurements will be used in the implicit version of the experiment. Lastly, participants will be assigned a specific order to progress through the experiments, chosen to mitigate potential biases such as perceiving the first experiment as the most challenging due to initial contact with the project.

The study will commence with a preliminary background form to gather information from participants, which may be relevant to draw some conclusions. Afterward, the participant will engage in one of three experiments and respond to the GEQ core, post and DDA questionnaire. After completing the three experiments, the participants will answer a series of questionnaires to identify their feelings toward the VR aspect of the project, its usability, and their preference between experiments (Fig. 24). Every session of the projected required approximately one hour, subject to slight variations, due to equipment malfunction.

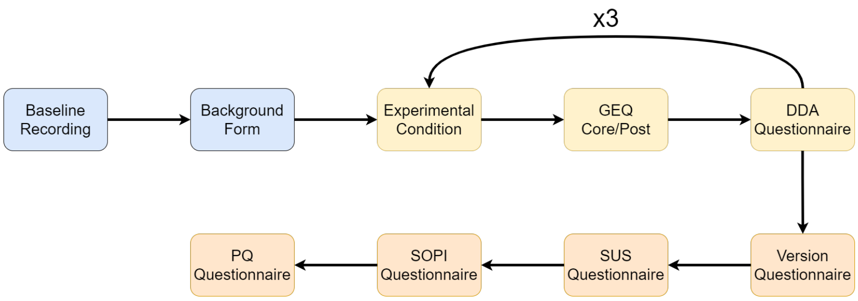


Fig. 24: Experimental Procedure

4.3.3 Setup

To put this progress in motion, we will need the proper setup to collect the maximum data possible and help avoid any misstep that will cause a breach in the fidelity of the data collected, putting the project at risk. So, to achieve the mentioned tasks, we will need a set of three computers, one HTC VIVE PRO EYE, a Polar H10 Heart Rate Sensor, the emteqPro and an exercise bicycle. Each one of these items has its crucial role to play, which are:

- **Cave-PC:** Used for resource-intensive programs, notably Robot Mania, a virtual reality (VR) game, and the Emteq Client, responsible for gathering physiological data during the experiment (Fig. 25a).This desktop was equipped with Windows 10,an AMD Ryzen 7 7700 3.80 GHz processor, 32 GB of RAM and an NVIDIA Quadro P6000 graphics card.
- **Desktop:**Tasked with running the main Pyflow program, which adapts the difficulty and records supplementary data of the PID Controller (Fig. 25b).This desktop was equipped with Windows 10, an Intel i7-6700 3.40 GHz processor with 16 GB of RAM.

- **Laptop:** Responsible for running a Pyflow program that records vital supplementary data (Fig. 25b).This desktop was equipped with Windows 10, an AMD Ryzen 5 3550H 2.10 GHz processor with 8 GB of RAM and an AMD Radeon Vega 8 graphics card.
- **emteqPro:** This equipment helps non-intrusively monitor heart rate during the experiment (Fig. 25c).
- **HTC VIVE PRO EYE:** VR headset for the project, easing player interaction with the game through its eye-tracking functionality and compatibility with the emteqPro sensor (Fig. 25d).
- **Polar H10 Heart Rate Sensor:** At the start, volunteers will wear the Polar H10 for one minute to establish a baseline heart rate reading. This baseline will be used as a reference point in one of the experiments (Fig. 25f).
- **Exercise Bicycle:** Provided to allow volunteers to impact heart rate directly during the experiment(Fig. 25e).

Another crucial detail to note is the varied setup for participants in each game. During the implicit version of the project due to the need to alter heart rate, participants will play the game while positioned on an exercise bike(Fig. 25h). For all other versions, players will be seated in a chair(Fig. 25g), as there is no requirement to vary heart rate, and more comfortable for the participant.

4.3.4 Participants

For this study, 26 participants were recruited from a convenience sample of volunteer subjects, all university students from the University of Madeira. This sample comprised individuals aged between 18 and 39 years old, with a sex balance of 14 females and 12 males. One participant was removed from the study due to equipment malfunction during data collection. An informed consent statement was provided before collecting data.

In conjunction with demographic data, some information about the gaming background was asked, as it may affect the volunteers' perspective. Some of the most relevant information gathered were:

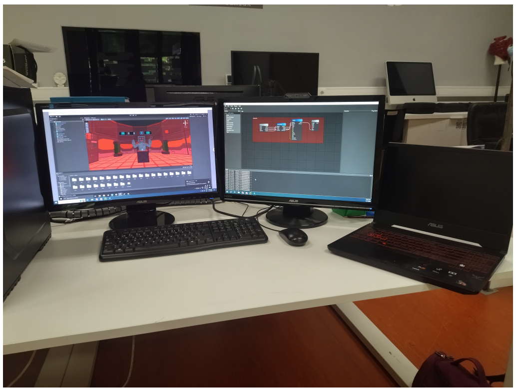
- **Preferred Difficulty:** The preferred difficulty was "Normal" (56%) followed by "Easy"(28%) and at last "Hard"(16%).
- **Difficulty Attempt:** From a scale of one to five in which one represents "never" and five "always" where the volunteers were asked how frequently they would play at certain difficulty levels. The results indicated that "Normal" (Median = 5) came out with the highest result, followed by "Hard" (Median= 3), and at last, "Easy"(Median= 2).

4.3.5 Data Analysis

The data analysis will encompass descriptive statistics and comparative analyses to assess and identify the relevant values collected during this experiment. Given the ordinary nature of the values collected during the experiment and the need to compare the repeated measures of each stimulus, the Friedman test was chosen as the most fitting test for the data analysis.

After identifying significant differences between groups (rejecting the null hypothesis), a post hoc analysis was conducted using the Wilcoxon test to explore these differences through pairwise

(a) Cave PC



(b) Desktop & Laptop

(c) emteqPro

(d) HTC VIVE PRO

(e) Exercise bicycle

(f) Polar H10 Heart Rate Sensor [57]

(g) Participant Setup A

(h) Participant Setup B

Fig. 25: Setup

comparisons. This test calculates the values between versions of the experiment. Then, if the values calculated show values inferior to the Bonferroni correction threshold, it would confirm the presence of a significant disparity between the versions compared. Data analyses were performed using SPSS 29.0.0.0 (IBM,US).

4.4 Conclusion

This section explores all the work done to validate the project's success. This includes the development of two programs for testing the project, "Robot Mania" and "EmteqClient." The first will be the game where the project will be tested. In contrast, the second one will record physiological data during the experiments used to adjust the difficulty of the implicit version of the project.

We also explore the user study, which will be intrinsic to answering the research questions raised by the project. This study will analyze the results from 3 experiments of the project, each focusing on a different type of stimuli: explicit, subjective, and implicit, and the affiliated questionnaires. These questionnaire results will be a window to the volunteer's perspective on the game experience, difficulty adjustment, and UI/UX of the project.

This comprehensive study aims to obtain the most realistic and accurate results possible to understand the project's limitations and successes better.

5 Results

This section will show the result of the usability test to verify the viability of the project by checking if the result verifies that the project reaches the objectives set at the beginning of the thesis. To confirm the project's viability, it must be demonstrated that the system consistently adjusts the game's difficulty to create an enjoyable, fair, and engaging experience and that the system changes the difficulty based on different stimuli in a way that the volunteers can consistently achieve the most optimal values. These points will be proven by questionnaires or the data collected during each experiment.

Before the data started, some errors in the collection should be noted. The volunteer 15 only completed their first experiment, the implicit experiment and the connected questionnaires. Furthermore, volunteer 10 did not respond to the GEQ-Core of the implicit experiment due to miscalculation. Finally in the Game Experience Questionnaire was found a lacking a question which could impact the evaluation of the aspect of challenge.

5.1 UI/UX

5.1.1 System Usability Scale Questionnaire

Volunteers express their feelings regarding the system's usability in a series of questions valued from 1 to 5, the first corresponding to "Strongly disagree" and 5 to "Strongly agree". Ratings were adjusted by subtracting 5 from even-numbered items and subtracting 1 from odd-numbered items. These scores were then summed, multiplied by 2.5 and the average of all the results was then compared to a scale from 0 to 100 to determine the system's usability level. In this dissertation, due to an oversight error, the scale was incorrectly valued from 1 to 7, resulting in the extra step of normalizing the values for a 5-point scale. The average of the study resulted in a value of 84.64 with a standard deviation of 10.85.

5.2 VR Experience

5.2.1 Sense of Presence Inventory Questionnaire

Concerning the results system of presence inventory as depicted in the box plot graph (Fig.26) below, the data was achieved by evaluating the answers of each volunteer by four different attributes, namely spatial presence, engagement, naturalness and Negative effects. Each attribute would have its value assigned by summing responses to specific questions. After each attribute's assessment for every volunteer, specific statistical values, namely the mean, and standard deviation (Table 11).

Table 11: SOPI Table.

	Mean	**SD**
Spatial Presence	2.76	0.78
Engagement	2.86	0.89
Naturalness	2.29	0.78
Negative Effects	1.63	0.60

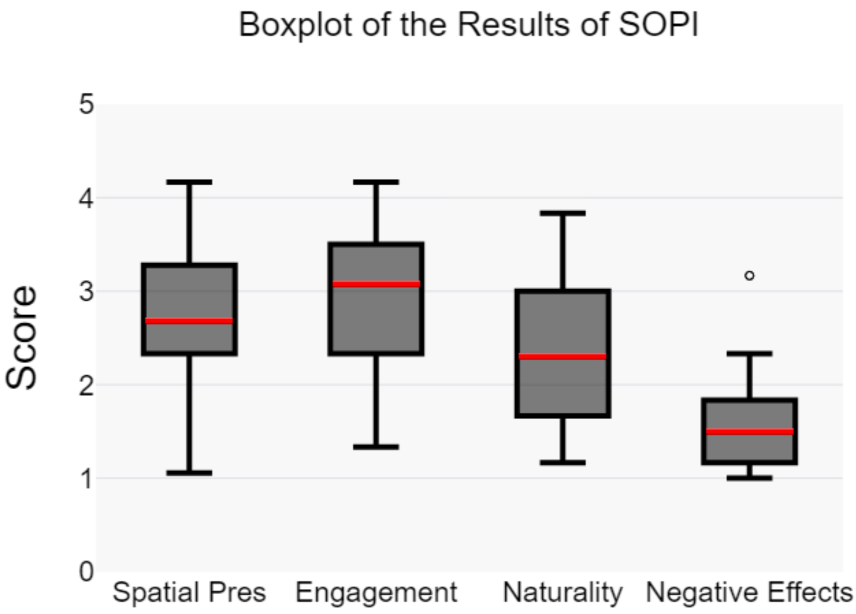


Fig. 26: Sense of Presence Inventory.

5.2.2 Presence Questionnaire

In this questionnaire, the volunteers expressed their feelings about their sense of presence inside the VR game in a series of questions using a 0 to 7-point scale." Each volunteer data contained volunteers' evaluations of realism, possibility to act, quality of the interface, possibility to examine, and self-evaluation of the performance felt by the volunteer. Organize each attribute's data in groups, and for each one, calculate the standard deviation, mean (Table 12). These values were utilized to generate a comprehensive box plot graph (Fig. 27) representing the questionnaire results.

Table 12: PQ Table.

	Mean	**SD**
Realism	5.17	0.86
Possibility to act	5.24	0.76
Quality of interface	4.24	1.16
Possibility to examine	4.57	1.07
Self Evaluation	5.44	1.27

5.3 Explicit

In this version of the experiment, the FOF uses the game score as an explicit real-time measure of the volunteers skill and success. The results gathered during this experiment were:

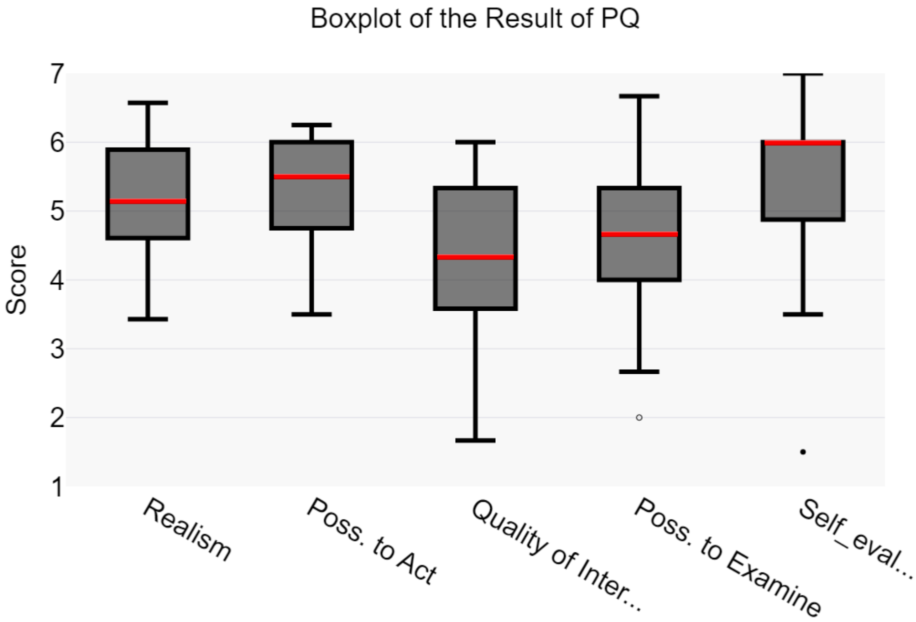


Fig. 27: Presence Questionnaire.

5.3.1 Game Data

The results of this version of the game show that, on average, the volunteers had achieved a score of 280.56 with a standard deviation of 37.25 at the end of the game.

In contrast, on average, the volunteers failed to collect 76.32 balls, with a standard deviation of 37.85 at the end of the game.

In total, it was estimated that volunteers, on average, caught 78.74% of the balls of the game with a standard deviation of 10.12.

5.3.2 Game Experience

In these results, we look to understand the experience of the volunteers during all versions of the experiment. By measuring the attributes, competence, immersion, flow, tension, challenge, negative affect, and positive affect. For the explicit version, the results were the following:

Table 13: Explict GEQ-Core

	Mean	**SD**
Competence	2.48	0.86
Immersion	1.87	0.59
Flow	2.42	0.90
Tension	1.04	1.12
Challenge	2.22	0.83
Negative affect	0.98	0.95
Positive affect	2.64	0.63

Based on the table. 13 results show that the explicit version has demonstrated that the volunteers had a higher approval for the dimensions of competence and positive affect (Mean =

2.48, SD=0.86). Followed closely by the flow (Mean = 2.42, SD=0.90), challenge (Mean = 2.22, SD=0.83), immersion (Mean = 1.87, SD=0.59), tension (Mean = 1.04, SD=1.12) and negative affect (Mean = 0.98, SD=0.95).

Table 14: Explict GEQ-Post

	Mean	SD
Positive Exp	2.07	0.83
Negative Exp	0.49	0.46
Tiredness	0.84	0.70
Returning to Reality	0.79	0.52

Analyzing the findings provided by the table. 14, which was created from the results of the GEQ Post game, which seek to understand the feelings of the volunteers after the explicit experience.The findings indicate that the feelings that were more prominent after the experience were of a positive experience (Mean = 2.07, SD=0.83). Sequentially, tiredness (Mean = 0.84, SD=0.70),return to reality (Mean = 0.79, SD=0.52), and negative experience (Mean =0.5, SD=0.46).

5.3.3 Adjustment of Difficulty

Employing explicit stimuli with FOF, as demonstrated through the results of the DDA questionnaire, on average, perceived that the difficulty matched their skill (Mean = 2.56, SD=0.96) rather than too elevated (Mean = 1.72, SD=1.43) or too soft (Mean = 1.52, SD = 1.42).

Concerning the adjustment of the difficulty, the volunteers felt more inclined to agree with being natural and organic (Mean = 2.32, SD = 1.18) than being significantly delayed (Mean = 1.24, SD = 1.09) or excessively rapid (Mean = 1.84, SD = 1.34).

Regarding the impact of the DDA system on the volunteers' experience, it was observed that the volunteers agreed that the system engaged (Mean = 2.56, SD = 0.96) substantially more as opposed to causing disruption (Mean = 1.00, SD = 1.22).

5.4 Subjective

This experiment variant involves the FOF utilizing the personal feeling to measure the volunteers performance. These measurements will be crucial to adapt the difficulty to the ideal level. The results gathered during this experiment were:

5.4.1 Game Data

From the results, it can be observed that, on average, the volunteers scored 326.08 with a standard deviation of 32.05 at the end of the game. The range of scores goes from a minimum of 281 to 390.

In contrast, the volunteers failed to collect, on average 35.56 balls, with a standard deviation of 26.66, at the end of the game. The number of balls that failed to collect had a minimum of 0 and a maximum of 89.

In total, it was estimated that the volunteers on average, caught 90.21% of the balls of the game with a standard deviation of 7.2.

Table 15: Subjective GEQ-Core

	Mean	SD
Competence	2.89	0.65
Immersion	1.95	0.78
Flow	2.54	0.87
Tension	0.55	0.59
Challenge	1.92	0.59
Negative affect	0.74	0.62
Positive affect	2.85	0.76

5.4.2 Game Experience

Based on the table. 15 results show that the subjective experiment has demonstrated that the volunteers had a higher approval for the dimensions of competence (Mean = 2.89, SD=0.65). Followed closely by positive affect (Mean = 2.85, SD= 0.76) , flow (Mean = 2.54, SD = 0.87), immersion (Mean = 1.95, SD = 0.78), challenge (Mean = 1.92, SD=0.59), negative affect (Mean = 0.74, SD=0.62) and tension (Mean = 0.55, SD=0.59).

Table 16: Subjective GEQ-Post

	Mean	SD
Positive Exp	2.17	0.93
Negative Exp	0.39	0.39
Tiredness	0.69	0.79
Return to reality	0.86	0.68

Analyzing the findings from the GEQ Post game (Table. 16), it can be attested that the feelings that were more prominent after the experience were a positive experience (Mean = 2.17, SD=0.93). Sequentially, return to reality (Mean = 0.86, SD = 0.68), tiredness (Mean = 0.69, SD=0.79), and negative experience (Mean = 0.39, SD=0.39).

5.4.3 Adjustment of Difficulty

Employing subjective stimuli with FOF, as demonstrated through the results of the DDA questionnaire, on average, perceived that the difficulty matched their skill (Mean = 2.80, SD = 0.87) rather than too advanced (Mean = 1.20, SD = 1.00) or too soft (Mean = 1.64, SD = 1.25).

Concerning the adjustment of the difficulty, the volunteers felt more inclined to agree with being natural and organic (Mean = 2.72, SD = 1.02) than being significantly delayed (Mean = 0.88, SD = 0.83) or excessively rapid (Mean = 1.40, SD = 1.08).

Regarding the impact of the DDA system on the volunteers' experience, it was observed that the volunteers agreed that the system engaged (Mean = 2.92, SD = 0.91) substantially more as opposed to causing disruption (Mean = 0.88, SD = 1.09).

5.5 Implicit

In this version of the experiment, the heart rate is used as a subtle indicator of the performance of the volunteer, to indicate the success and skill of the individual. These values will be used to tailor the ideal challenge for the volunteers. The results gathered during this experiment were:

5.5.1 Game Data

From the results, it can be observed that, on average, the volunteers scored 378.04 with a standard deviation of 47.39 at the end of the game. The range of scores goes from a minimum of 313 to 508.

In contrast, the volunteers failed to collect, on average, 7.24 balls, with a standard deviation of 12.76, at the end of the game. The number of balls that failed to collect had a minimum of 0 and a maximum of 51.

In total, it was estimated that the volunteers, on average, caught 98.23% of the balls of the game with a standard deviation of 3.21 .

5.5.2 Game Experience

Table 17: Implicit GEQ-Core

	Mean	**SD**
Competence	2.91	0.75
Immersion	1.84	0.78
Flow	2.38	1.09
Tension	0.54	0.92
Challenge	1.33	0.74
Negative affect	0.86	0.80
Positive affect	2.62	0.97

Based on the table. 18 results show that the implicit experiment has demonstrated that the volunteers had a higher approval for the dimensions of competence (Mean = 2.91, SD=0.75). Followed closely by positive affect (Mean = 2.62, SD=0.97), flow (Mean = 2.38, SD=1.09), immersion (Mean = 1.84, SD=0.78), challenge (Mean = 1.33, SD=0.74), negative affect (Mean = 0.86, SD=0.80) and tension (Mean = 0.54, SD=0.92).

Table 18: Implicit GEQ-Post

	Mean	**SD**
Positive Exp	2.40	1.11
Negative Exp	0.48	0.57
Tiredness	0.84	0.77
Returning to Reality	0.89	0.61

The findings from GEQ Post game (Table. 18) indicate that the feelings that were more prominent after the experience were of a positive experience (Mean = 2.40, SD=1.11). Sequentially, return to reality (Mean = 0.89, SD=0.61), tiredness (Mean = 0.84, SD=0.77), and negative experience (Mean = 0.48, SD=0.57).

5.5.3 Adjustment of Difficulty

Employing implicit stimuli with FOF, as demonstrated through the results of the DDA questionnaire, on average, perceived that the difficulty matched their skill (Mean = 2.88, SD = 0.78) rather than too elevated (Mean = 0.84, SD = 0.99) or too soft (Mean = 2.36, SD = 1.08).

Concerning the adjustment of the difficulty, the volunteers felt more inclined to agree with being natural and organic (Mean = 2.84, SD = 0.99) than being significantly delayed (Mean = 1.68, SD = 1.35) or excessively rapid (Mean = 0.32, SD = 0.56).

Regarding the impact of the DDA system on the volunteers' experience, it was observed that the volunteers agreed that the system engaged (Mean = 2.72, SD = 1.24) substantially more as opposed to causing disruption (Mean = 0.40, SD = 0.82).

5.6 Statistical Analysis

This section of the result will explain the results of the statistical analysis done in the project in intricate detail. For a start, we used the data collected from the attributes of the questionnaires GEQ Core and Post-game and the score of the volunteer for each experiment to compare and discover any significant difference between experiments. This analysis revealed substantial changes, showing significant change between the three attributes:

5.6.1 Tension/Annoyance

The statistical analysis reveals that volunteers notably perceive with a significant contrast the attribute Tension / Annoyance (Fr(2) = 6.19, p=0.045) between versions of the Stimuli. Following post hoc analysis employing Wilcoxon signed-rank tests and Bonferroni correction, the results show that no significant differences were observed between the tension felt in each experiment.

5.6.2 Challenge

The statistical analysis reveals that volunteers notably perceive with a significant contrast the attribute Challenge (Fr(2) = 15.371, p<0.001) between versions of the Stimuli. Following post hoc analysis employing Wilcoxon signed-rank tests and Bonferroni correction, the results reveal (Fig. 31a) that volunteers notably perceived a significantly richer challenge in the Explicit (Z = -3.724, $p = 0.001$, two-tailed) and Subjective (Z = -2.906, $p = 0.004$, two-tailed) versions compared to the Implicit version.

5.6.3 In-Game Score

The statistical analysis reveals that volunteers notably perceive with a significant contrast the attribute In-Game Score (Fr(2) = 34.16, p<0.001) between versions of the Stimuli. Following post hoc analysis employing Wilcoxon signed-rank tests and Bonferroni correction, the results show (Fig. 28) a noteworthy discrepancy between the in-game scores from all the versions of the experiment, with the scores being higher in the Implicit version compared to the Explicit (Z = -4.373, $p < 0.001$, two-tailed) and Subjective (Z = -3.350, $p < 0.001$, two-tailed) versions.Upon further analysis, it is evident that a significant disparity exists between the Explicit and Subjective versions(Z = -3.835, $p < 0.001$, two-tailed) versions, with the Subjective score surpassing that of the Explicit.

5.7 System Overall

This section aims to provide a comprehensive of the three experiments realized in this dissertation. To be more specific, compare results from the game, the volunteers' feelings during and after the experience, and their opinions towards the adjustment difficulty.

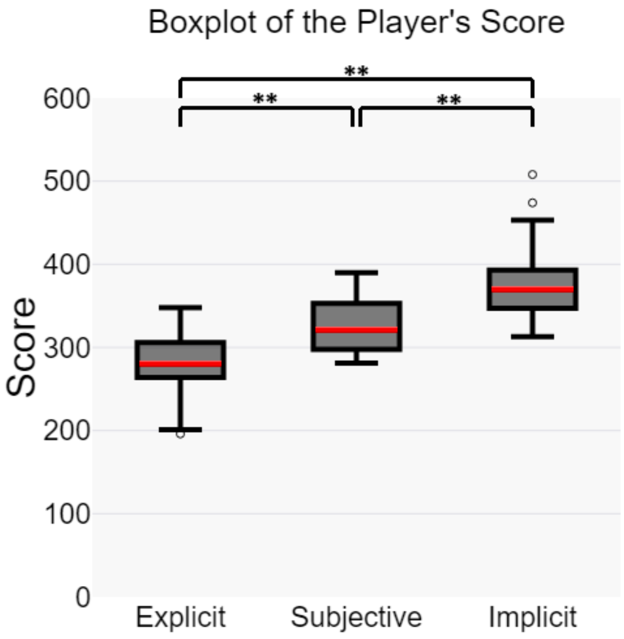


Fig. 28: In-Game Score Results **p-value < 0.001.

5.7.1 Game Data

To understand how the experiments affect the volunteer's performance in the game, it saves data from different aspects, the most relevant being the volunteer's score during the game. Looking at the average result of the volunteers in each experiment, it becomes evident that the implicit version shows a higher on-average score, followed by the subjective version of the project and the last explicit version.

5.7.2 Game Experience

Looking at the feeling towards the GEQ core results (Fig.29), it can be determined that the feeling of competence (Fig.30a) were higher during both the implicit (Mean = 2.91, SD = 0.75) and subjective (Mean = 2.89, SD = 0.65) compared to the explicit (Mean = 2.48, SD = 0.86) version. In terms of Immersion (Fig.30b), it can be deduced that the subjective (Mean = 1.95, SD = 0.78) experiment was ranked highest in terms of immersion, followed by the explicit (Mean = 1.84, SD = 0.78) and finally the implicit (Mean = 1.84, SD = 0.78).

The experimental flow, as illustrated in Figure 30c, exhibited a higher mean score in the subjective experiments (Mean = 2.42, SD = 0.90) and implicit (Mean = 2.38, SD = 1.09) experiments. In comparison, the feeling of tension and annoyance (Fig.30d) can be observed to be higher in the explicit (Mean = 1.04, SD = 1.12) experiment, followed by the subjective (Mean = 0.55, SD = 0.59) and implicit (Mean = 0.54) experiment. In the dimension of challenge (Fig.31a) the explicit (Mean = 2.22, SD = 0.83) experiment was first followed closely by the subjective (Mean = 1.92, SD = 0.59) and implicit (Mean = 1.33, SD = 0.74) experiments.

Overall, the rated low in terms of negative feeling (Fig.31b), the subjective (Mean = 0.74, SD = 0.62) version being the lowest and followed by the implicit (Mean = 0.86, SD = 0.80) and explicit

(Mean = 0.98, SD = 0.95). In contrast, the subjective version (Mean = 2.85, SD = 0.76) generated the most positive experience(Fig.31c), compared to the explicit (Mean = 2.64, SD = 0.63) and implicit (Mean = 2.62, SD = 0.97).

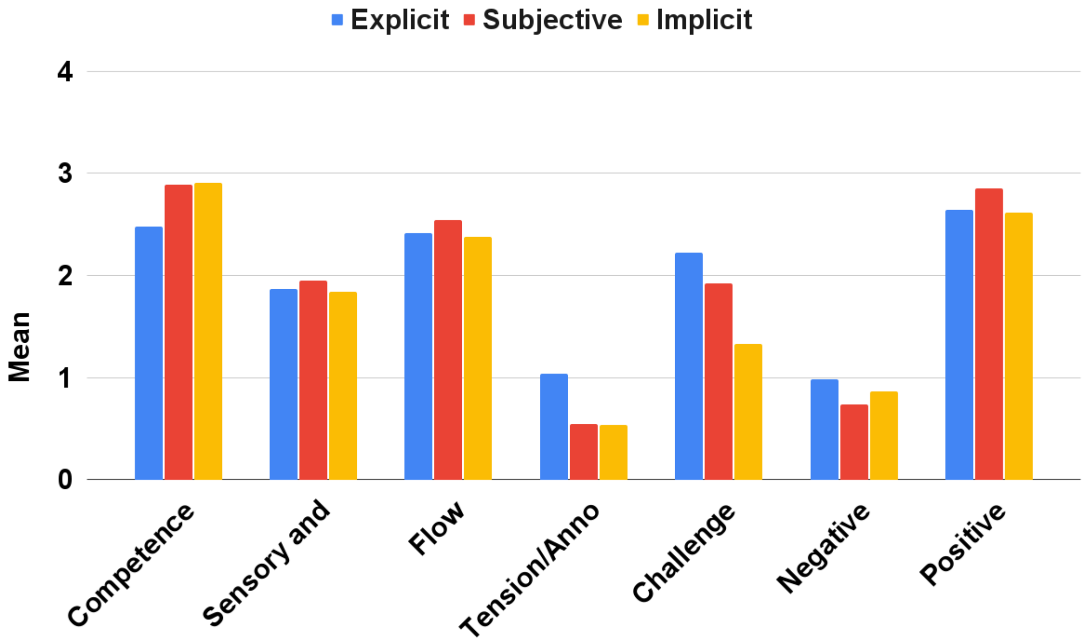


Fig. 29: GEQ Core Results.

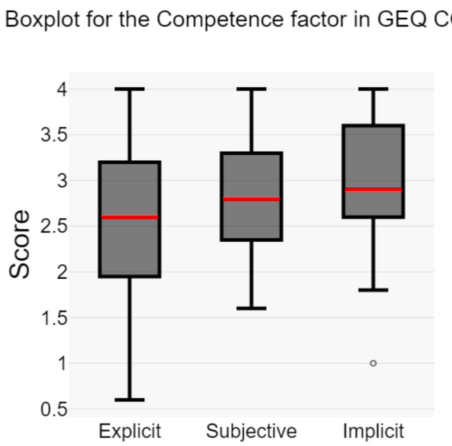
Looking at the opinion of the volunteers about their feelings in the aftermath of each experiment by using the GEQ Post (Fig.32), it can be observed that a higher positive experience (Fig.33a), in implicit experiment (Mean = 2.41, SD = 1.11), followed by subjective(Mean = 2.17, SD = 0.93) and explicit(Mean = 2.07, SD = 0.83). In contrast, it can be seen a superior level of negative experience (Fig.33b) in the explicit (Mean = 0.49, SD = 0.46) version, lower it can be found the implicit(Mean = 0.48, SD = 0.57) and subjective(Mean = 0.39, SD = 0.39) experiment.

After the experience, the volunteers reported a higher level of tiredness (Fig.33c) after the explicit (Mean = 0.84, SD = 0.77) and implicit (Mean = 0.84, SD = 0.70) experiments feeling less tired in the subjective (Mean = 0.69, SD = 0.79) versions. The return to reality (Fig.33d) was lower in the explicit (Mean = 0.79, SD = 0.52) version compared to the implicit(Mean = 0.89, SD = 0.61) and subjective (Mean = 0.86, SD = 0.68) version.

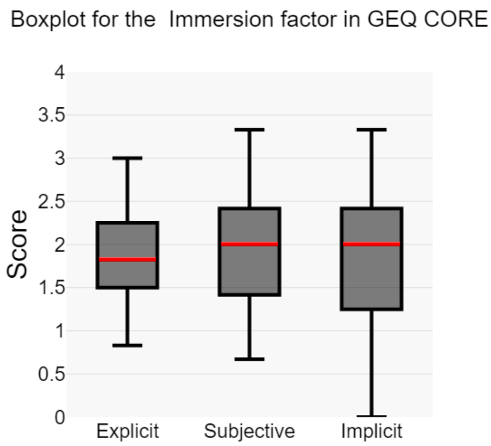
5.7.3 Adjustment of Difficulty

Looking at the diverse results from the DDA Questionnaire, results can be shown to be comparable, though with some deviations. The results identified across all the experiments, the difficulty was considered a match (Fig.34a). Although the explicit has shown the highest difficulty values (Mean = 1.72, SD = 1.43), the implicit has shown the highest easiness (2.36, SD = 1.08) and matching difficulty (Mean = 2.88, SD = 0.78).

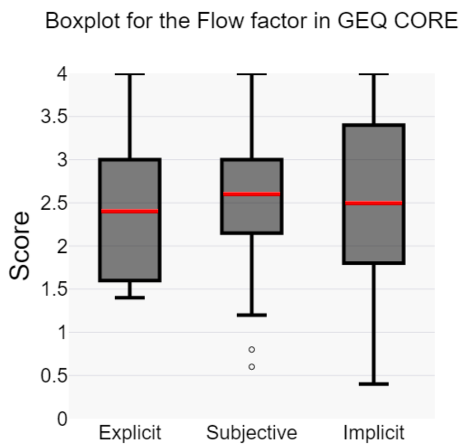
In terms of the adjustment of difficulty overall, the highest values across experiments were "Natural & Organic" with some deviation (Fig.34b). It can also be seen that the explicit version



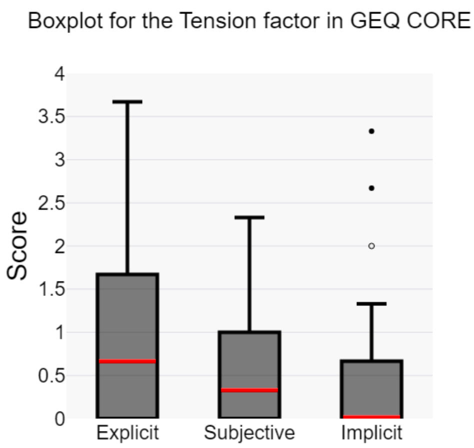
(a) Competence



(b) Sensory and Imagination

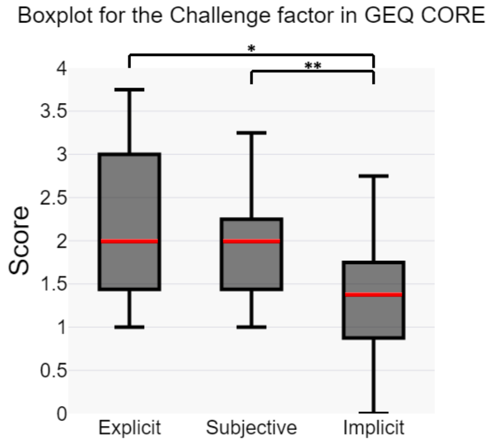


(c) Flow



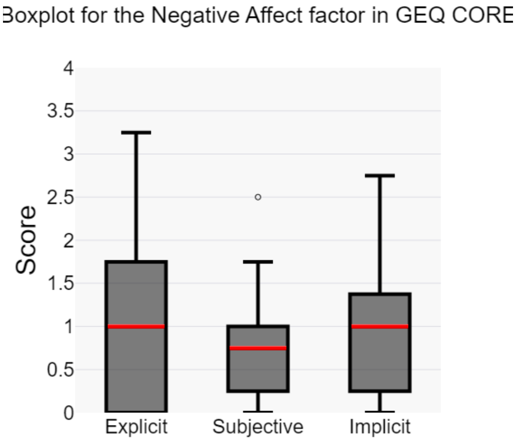
(d) Tension Annoyance

Fig. 30: Box Plot GEQ-Core Results - Page 1

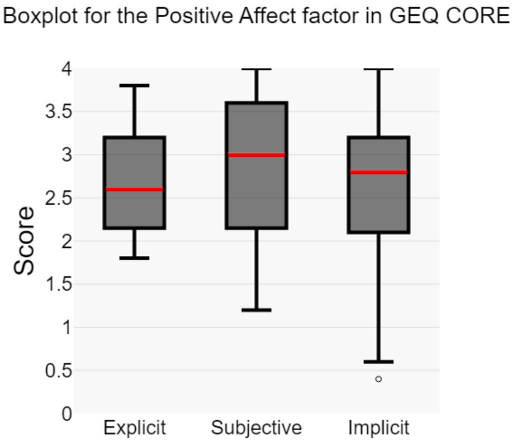


*p-value < 0.017, **p-value < 0.001.

(a) Challenge



(b) Negative Affect



(c) Positive Affect

Fig. 31: Box Plot GEQ-Core Results - Page 2

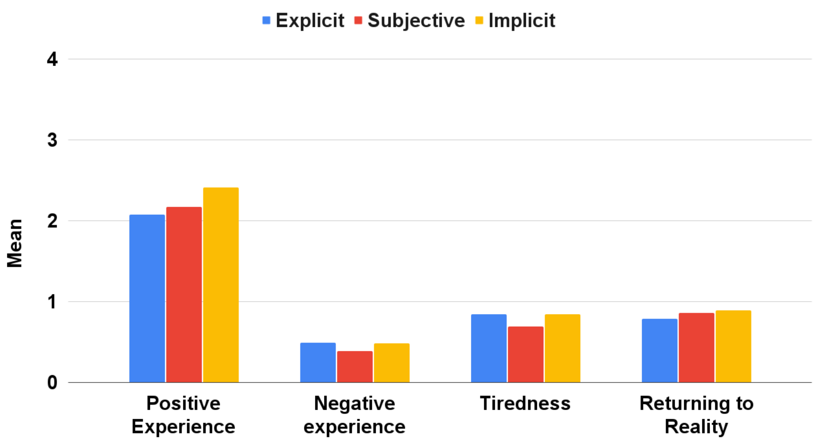
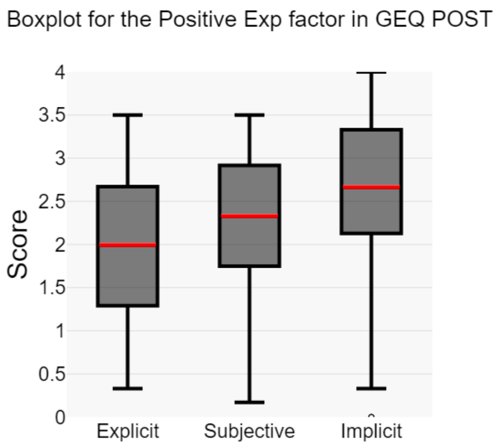
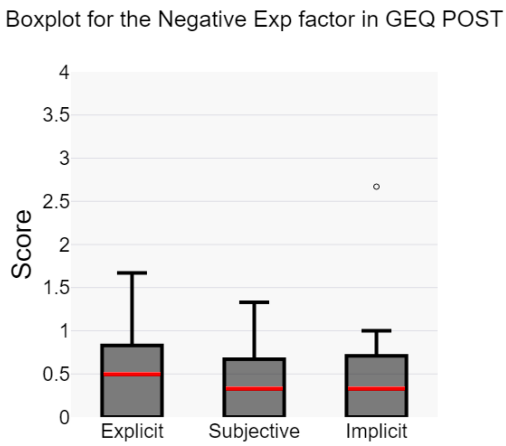


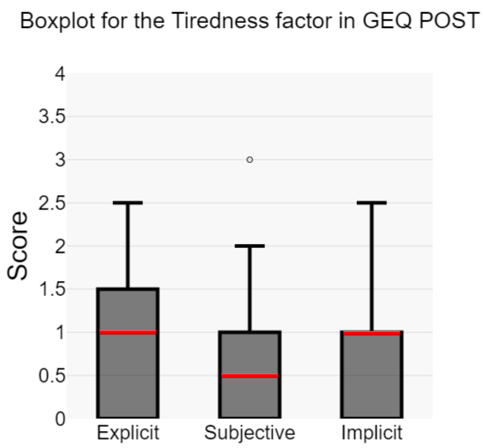
Fig. 32: GEQ Post Results.



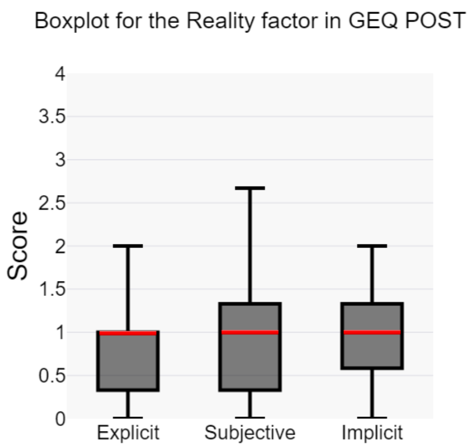
(a) Positive Experience



(b) Negative Experience



(c) Tiredness

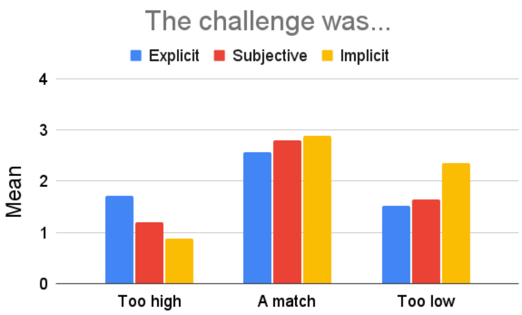


(d) Returning to Reality

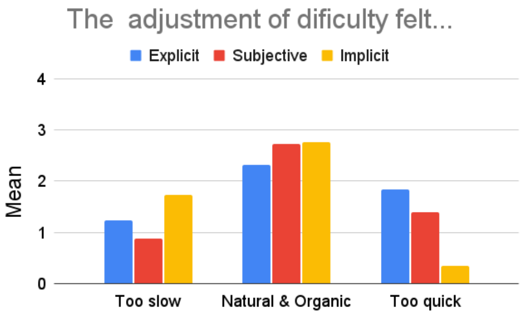
Fig. 33: Box Plot GEQ-Post

shown to suffer the most from a too quick adjustment of difficulty (Mean = 1.84, SD = 1.34). In contrast, the implicit version is shown to be the most organic (Mean = 2.84, SD = 0.99) but also voted the slowest (Mean = 1.68, SD = 1.35).

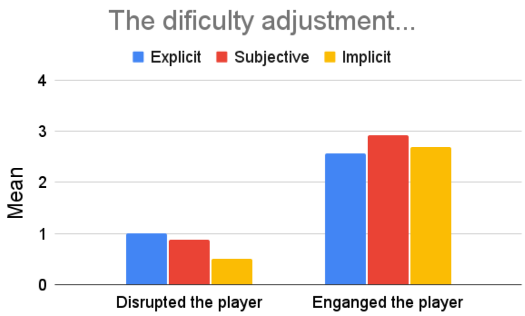
The influence of DDA system across the experiments was seen to engage (Fig.34c) the volunteers with some small deviation in the values. The explicit version (Mean = 1.00, SD = 1.22) has been shown to be the most disruptive and the subjective (Mean = 2.92, SD = 0.91) the most engaging.



(a) Challenge



(b) Adjustment of Difficulty



(c) Experience created

Fig. 34: DDA Questionnaire Graph

5.7.4 Version Comparison

Within this questionnaire, the volunteers reply to several questions where they must choose a version over others to determine their overall feeling towards each version. The collection of the volunteers' answers (Fig.35) has revealed that in general volunteers, in terms of enjoyment, the explicit response registers a 28%. At the same time, the subjective and implicit experiments present us with 24% and 48%, respectively.

Regarding the perceived challenge, the explicit version indicates 56% while the subjective records 8% and the implicit 36%.On the recommendation topic, the explicit was endorsed at 16%, with subjective and implicit endorsements at 32% and 52%, respectively.

Regarding difficulty in the experiment, the explicit is dominant at 56%, while subjective assessments record 12%, and implicit stands at 32%. By contrast, the subjective experiment is perceived as the easiest by 48% juxtaposed against the explicit and implicit with 20% and 32%, respectively.

The experiment had higher endorsement about matching skill with difficulty was implicit with 44%, followed by the subjective with 36% and the last explicit version with 20%. In contraposition, the explicit version with 44% was voted as matching less the volunteers' skill while the implicit followed with 28% and the subjective with 24%.

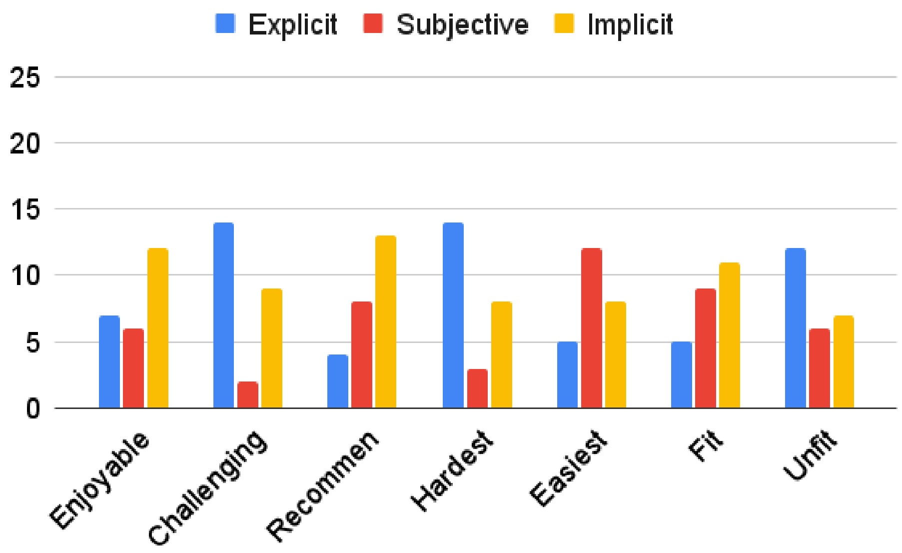


Fig. 35: Version Questionnaire.

5.8 Conclusion

This section has shown the results of the user study performed for this dissertation. The methodology employed, where the volunteers will test a series of questionnaires (e.g., Game Experience, SUS, SOPI) of three different experiments created to use different stimuli explicit, subjective, and implicit in adapting to the difficulty, allowed us to investigate if the framework can adapt to different stimuli and if it creates an enjoyable and engaging experience for all the stimuli.

From the findings, overall, the volunteers perceived every experiment's positive aspects, like competence, flow, and positive experience, with higher regard. In contrast, more unwanted aspects, like tension, negative affect, and tiredness, were perceived with lesser regard. It can also be noted that across the different experiments, the difficulty adjustment was rated higher as being a match for the volunteers' skill, the changes of difficulty being organic, natural, and engaging.

Regardless, these results alone do not entirely represent volunteers' perspectives. The following section is a comprehensive discussion of the meaning behind the results acquired during the experiment.

6 Discussion

In this thesis, we presented the FOF framework to enhance the state of flow in serious games using DDA. From the analysis of the diverse data collected, a reflection on the project success and failures. These reflections will then be used as guides for future possible studies.

6.1 UI/UX

UI and UX are crucial parts of any game, as they both enhance the interaction between the participant and the game and the system's usability.

6.1.1 System Usability Scale Questionnaire

To evaluate this critical aspect, the System Usability Scale was utilized, in which the system has to score above 68 to show the system has a good UI/UX. The results from the usability study revealed that our framework has a high perceived usability score (84.64 ± 10.85) compared to previous studies that evaluated the usability of their systems [54].

6.2 VR Experience

As important as the UI and UX, the VR experience is a crucial component of the experience of the project. A poorly designed VR experience can damage the focus and enjoyment of the experience [58] and undermine the progress made by DDA. In order to evaluate the VR experience from a user's perspective, two questionnaires were employed, namely, the SOPI and the PQ.

6.2.1 Sense of Presence Inventory Questionnaire

Findings from the SOPI questionnaire (Fig.26), revealed that the volunteers felt moderately present in the virtual environment as shown by the medium level of spatial presence, however, they did not perceive it as highly realistic, as evidenced by the low scores on naturalness. In contrast the volunteers felt involved in the game; as we can see in the graph (Fig.26), the volunteers experienced a higher sense of engagement while the negative effects were decreased. Overall, the volunteers felt that they experienced an engaging but not realistic experience.

6.2.2 Presence Questionnaire

Similarly, the analysis of data from the PQ questionnaire (Fig.27), it can be noted that all the characteristics evaluated can be pronounced positive due to scoring averagely above the middle-value of 4. This value indicates that overall, the volunteers felt immersed in the world created and had no major issues with the experience.

By combining these insights from the volunteers we concluded that ultimately, the VR experience successfully engaged them, However, the limited realism However, the limited realism.

6.3 Experiments comparison

Moving into further details, we evaluated different aspects of the experience between the three versions. From these aspects, we seek to understand the volunteers perspective on the experience to determine if FOF was able to adapt the difficulty of the game using different stimuli.

6.3.1 Game Data

Analyzing the in-game data retrieved from the experiments, conclusions about the concrete facts about the volunteers' performance in the game were arrived at, as it can be observed that, on average, volunteers played the game well, as can be noted by the low error rate found across the participants. Another conclusion was that the FOF created a similar challenge for the volunteers, as can be seen by the low standard deviation of the catch rate. From each experiment, specific details were noted:

- **Explicit:** This specific experiment was designed to maintain the players in a constant error rate of 50%, however, results show 21% error rate. This discrepancy can be attributed to the difficulty range imposed on the system, which is indicated by the low variability of the results. This suggests that if the maximum level of difficulty was higher, the error rate might have been higher.
- **Subjective:** The low error rate of 10% shown in this experiment might indicate a lack of challenge for the volunteers. This can be attributed to the focus of this experiment as, in contrast to its explicit counterpart, it focuses on maintaining the volunteer in a state between stress and relaxation instead of focusing on creating a challenge on the volunteer's skills.
- **Implicit:** The low error rate of 3% shown in this experiment might indicate a lack of challenge for the volunteers. This can be attributed to creating challenges indirectly by obscuring the screen and the exercise needed to maintain the heart in the experiment threshold.

To reiterate from the concrete data gathered during all the experiments, it can be concluded that the FOF dynamic difficulty had a similar impact across all the volunteers population in each game. Which includes the lack of challenges faced across the different experiments.

6.3.2 Game Experience

From the data from GEQ Cores questionnaires of each experiment, some insightful information can be extrapolated about the perceived experience of the volunteers during the experiments:

- During the experiment, volunteers perceived a moderated level of competence. This result leads to the suggestion that the game is not too demanding or too comfortable.
- The immersion scores suggest volunteers may not have been fully engaged with the game. This could be due to the technical limitations that disturb the immersion of the volunteer or due to the repetitive nature of the game. Despite this, the flow experienced by volunteers suggests a focused state during the gameplay despite not reaching optimal levels.
- The low tension value indicates that the participants generally did not experience stress or anxiety during the projects. Also noteworthy is the low to moderate level of challenge, which may indicate the FOF was able to adjust difficulty correctly, as was noted from the background form of the volunteers a 56% as a preference for a more medium level of challenge.
- Additionally, the low negative affect values imply that volunteers suffered no significant negative impact. In contrast, the moderately high positive affect indicates that the experiments positively influenced the volunteers.

Regarding the GEQ Post Game results, we can draw some conclusions about the feelings of the volunteers following their participation in all the experiments.

- The experiment elicited a moderately high level of positive experience, implying that it was generally appreciated and found enjoyable by the majority of participants. This can also be confirmed by the low level of negative experience, indicating a low level of adverse reaction.
- The level of tiredness and returning to reality being low indicates that the experiment had minimal impact on participants' energy levels, and the transition from game to reality was not overly challenging for most volunteers.

To reiterate, from these results, it can be observed that during and after all the experience, the volunteers considered the game experience as positive and engaging, making them feel competent while avoiding fatigue. On the other hand, the volunteers were not as immersed in the environment created and lacked a sufficient level of challenge.

6.3.3 Adjustment of Difficulty

Within this DDA questionnaire, we sought volunteers' perceptions of the implemented difficulty adjustment. The results of this questionnaire were divided into three key areas, aiming to assess whether the challenge matches the skill of the volunteers, "How did the volunteers feel about the adjustment?" and "Did the engagement of difficulty engage or disrupt the volunteer?".

Related to the first question, people were asked three times on a scale from 0 to 4 if during the game the challenge they felt was too high, too low, or a match. The results of this inquiry show that, in general, the volunteers felt the challenge corresponded well with their skill levels, as can be seen in Fig.34a.

Concerning the adjustment, volunteers were asked three questions: Was the difficulty adjustment too quick, too slow, or organic and natural? The results, portrayed in Fig.34b, reveal that the volunteers felt the changes were organic and natural, indicating that the system did not suffer from an inconsistent adaptation of difficulty.

Regarding the influence of difficulty adjustment on the volunteer experience, questions were posed to elicit responses, and one question asked if the adjustment was perceived as disruptive and the other if it engaged the volunteer. By the data extracted (Fig.34c), we can conclude that the volunteers expressed a more pronounced sentiment towards the adjustment being engaging rather than disruptive.

From all these results, the implementation of the DDA system in all the experiments was shown to avoid some of the most common pitfalls of implementing DDA, like difficulty balancing, lack of control of the difficulty, inconsistency and player frustration.

6.3.4 Version Comparison

In this questionnaire, we ask several questions about the different feelings between the three versions of the experiment. The results of this questionnaire show us that (Fig.35):

- **Explicit:** This version of the experiment is considered the most complex and challenging while being the one that most volunteers felt matched their in-game abilities. This feeling was justified by most of the volunteers by referring to the levels of speed of the game, making it hard to gather points.
- **Subjective:** This version of the experiment is most qualified as the least challenging version of this project while being tied with an implicit version for the experiment in which the difficulty

matches the difficulty the best. These volunteers generally justified these opinions by explaining that they felt that the difficulty adjustment was slower, giving more time to adapt.

- **Implicit:**This version of the project has been demonstrated to provide a more enjoyable experience compared to the other two experiments, and tied with the subjective version as the experiment that create the fairest experience. The volunteers justified this feeling by saying that the addiction to physical exercise added a layer of difficulty that made the game more enjoyable.

In summary, in all the experiments the volunteers overall showcased a high aptitude for the game, although with low variations in results, suggesting a relatively equal impact of difficulty adjustments on volunteers. The GEQ Core implied a nuanced picture of moderated feeling competence and flow, in conjunction with low tension, medium-low immersion and challenge, and a positive affect during the experiment. Post-game evaluation showcases positive experience with minimal fatigue. The DDA questionnaire underscores the volunteers perception of adjustment of difficulty created by the FOF as a match to their skills, natural and engaging. These findings suggest that while the experiment was engaging and appropriately matched the volunteers' skills, it was perceived as lacking in challenge, as reflected in the participants' high success rates and the moderate GEQ core scores related to challenge.

Across all the comparisons done in this section, the explicit experiment was considered consistently the most challenging. This was primarily attributed to the high speed at which the balls could approach and the focus in this experiment on maintaining a challenge for the volunteers. This data aligns with the data collected from the game and GEQ Core, DDA and Version questionnaire.

In most cases, the subjective version was found to be a middle ground between the implicit and explicit counterparts. Key factors such as immersion, positive affect during the experiment, and engagement created by the DDA system highlighted it as superior. One thing to note is that it is generally considered a middle value between implicit and explicit in terms of challenge and difficulty, as collaborated by the GEQ Core, DDA questionnaires and in-game score. However, when the participant was asked to rate them, it was considered the easiest and the least challenging. This version's simplicity and slower adjustment might justify this contradiction, making it feel like a lesser challenge than other counterparts.

6.4 Statistical Analysis

The results of the Friedman test, followed by post hoc Wilcoxon signed-rank tests with Bonferroni correction, showed a more perceptive view of the participant's sense of tension/annoyance and challenge and the In-game score across the different experiment counterparts.

6.4.1 Tension/Annoyance

Although the Friedman test showed a significant contrast in volunteers' perception of Tension/Annoyance across different stimuli versions. The post hoc, on the other hand, did not find any significant difference between all experiments. This indicates a difference between the perceptions, although this discrepancy is not enough to be shown in a pairwise comparison.

6.4.2 Challenge

The analysis of the Challenge attribute revealed a considerable contrast among versions of the stimuli. The post-hoc test revealed a pronounced difference between the experiments, to be precise, between the explicit and implicit and subjective and implicit pairs, indicating there is no pronounced difference between the explicit and subjective pair. These results indicate that the implicit version lacks challenge compared with the other two counterparts.

6.4.3 In-Game Score

Volunteers' perception of the In-Game Score feature displayed a significant disparity among different renditions of the experiment. It was observed from the post hoc analysis that between every stimulus, the most prominent was the explicit and implicit. These prominent contrasts indicate a substantial difference in how the stimulus influences the participants' performance.

This analysis suggests that various stimuli had distinct impacts on participants' experiences and performances. While a significant contrast was noted between the perception of tension/annoyance when compared to pairwise, no specific difference was noted, indicating that volunteers did not perceive the tension between the experiments in a way that shows a clear distinction. Another thing to note is the dichotomy of the lack of significant difference between the subjective and explicit experiments' challenges, and the significant difference between the scores implied a disconnect between the perspective of the volunteers and the results from the game.

6.5 Implication

From the results achieved during this study, some of the implications are uncovered, from the perspective of the literature previously gathered. From the neurorehabilitation standpoint, the project followed certain key points to ensure motivation by creating an optimal experience for making the project viable for use in neurorehabilitation. These key points were followed as expressed in the previous sections (Section 3) and their effectiveness was confirmed by the results of the GEQ Core and Post as both questionnaires indicate a volunteer had an experience which was viewed as positive and engaging.

When juxtaposing the FOF GEQ results (Fig.29) with another DDA project, VR Exergames [1] (Fig. 36), a project created with the objective to motivate the participants to do exercise regularly thought the virtual reality, some interesting observation can be made:

Firstly, the value of competence in both studies showed similar results. In contrast, immersion, flow, and positivity have presented higher values in the Exergame study. These higher results can be caused by a myriad of reasons, from the focus on exercise to making the experience more engaging to using a more complex algorithm as reinforcement learning, creating a balanced and immersive experience when compared with a more effortless DDA algorithm like the PID controller.

From these observations, we can take away signs that will be valuable for any future work done on this project as they show the potential of factors that more complex algorithms might have in enhancing the experience and, together with the motivation of the experience, making this project fitter for neurorehabilitation.

From the perspective of the research done on the topic of DDA, FOF followed all the steps to maximize the success of the DDA implemented from the game, from carefully choosing facets

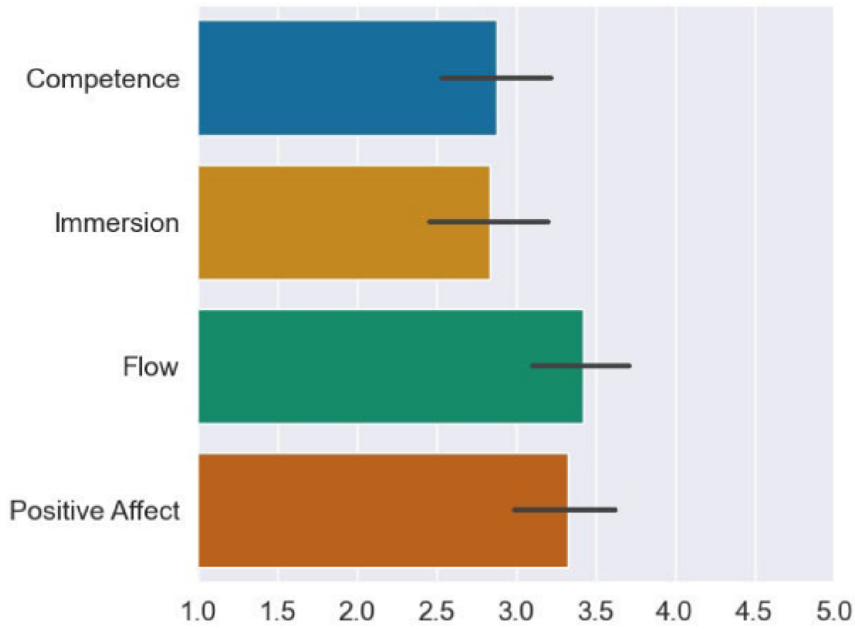


Fig. 36: VR Exergames GEQ Results [1]

of the process of DDA while also adding to the new solutions to solve some of the most recurring problems. The challenges faced in DDA, as identified in the state of the art, include Player feedback, Complexity, Testing and debugging, Integration with other game systems, Game-Specific Tailoring, Player Frustration, Lack of Control, Inconsistency, and Linearity [4, 11, 46–48]. Through a meticulous process, FOF managed to mitigate most of these challenges.

One of the challenges tackled successfully was an unbalanced difficulty, as indicated by the values obtained from the DDA questionnaire, which shows it can be observed from the overall volunteers' perspective that the difficulty matched their skills. Similarly, this questionnaire also showed the adjustment as natural and organic while creating an engaging experience that dispels the concerns over the adjustments' lack of control or inconsistency.

Following with another common DDA problem is integrating the system with other game components. However, the DDA framework's ability to use any Unity game script simplifies the use of variables from these complex components without disrupting their flow. Furthermore, when paired up with the PID Controller's ability to be used as a generic algorithm, the DDA framework makes the system adaptable to most games without much tailoring.

Another common problem that FOF tackled was the frustration that arises from a flawed DDA system. The results of the GEQ Core imply that FOF successfully addressed this issue, as they indicate an overall absence of significant levels of frustration.

One problem that FOF cannot avoid, but has the ability to alleviate, is the large amount of testing required to get the DDA system to a level that fits the needs of the user. Pyflow visual scripting features make this process easier and quicker. It streamlines the workflow and spares developers from adapting code and rebuilding the project, thereby increasing productivity. Additionally, FOF mitigates the complexity of using a DDA system, although not entirely, as complexity is an inherent quality. However, the system still encounters issues related to linearity because it adapts to player skill rather than preferences, which are inherently subjective.

This inability to cater to people with a preference a static difficulty is very common, as can be seen in the results from other similar projects as per Souza et al. [30]. This setback might be a focus of future research if the project continues out of the bounds of this thesis.

In summary, the FOF approach to DDA implementation addressed most of the common problems with DDA while also improving both the volunteer's experience and developer efficacy.

6.6 Experiment Reflection

Across all the experiments mentioned in this section, the explicit experiment was considered consistently the most challenging. This was primarily attributed to the high speed at which the balls could approach and the focus in this experiment on maintaining a challenge for the volunteers. This data goes in line with the data collected from the game and GEQ Core, DDA and Version questionnaire.

In most cases, the subjective version of the experiment served as a median between the implicit and explicit variants. Key factors such as immersion, positive affect during the experiment, and engagement created by the DDA system highlighted it as superior. One thing to note is that it is generally considered a middle value between implicit and explicit in terms of challenge and difficulty, as collaborated by the GEQ Core, DDA questionnaires and in-game score. However, when the participant was asked to rate them, it was considered the easiest and the least challenging. This version's simplicity and slower adjustment might justify this contradiction, making it feel like a lesser challenge than other counterparts.

6.7 Conclusion

Overall, based on the results obtained, the Flow Optimizer Framework adjusted the game difficulty so that the volunteers experienced an engaging and enjoyable experience while using different stimuli in the project as a source of performance. While some experiments showed positive outcomes, there are identified areas for improvement in the challenge and immersion aspects of the system.

7 Overall Conclusion

Within this section, we briefly and understandably describe the motivation, problems, solutions, limitations, and results found during the execution of this dissertation. We also discuss future options for continuing this dissertation.

7.1 Summary of Research

Initiating the discussion, we must explain all the processes, starting with the motivation that pushes us to follow through. Our core objective is to help rehabilitation become a more fun and engaging experience, making the patient happier and more focused on rehabilitation, ultimately enhancing rehabilitation outcomes. To achieve this, we aim to create an agnostic DDA system seamlessly applicable to a wide range of games, the framework Flow Optimizer.

The efficacy of this system relies on positive responses to our research questions. These are whether the system created can provide most people with an engaging and fun experience and exhibit the ability to adapt the difficulty using various stimuli.

Focusing on the first question, looking at the results from the GEQ Core and Post, we can conclude that the results were positive due to, on average, all experiments having high feelings of flow, competence, positive affect, and positive experience, and a low level of endorsement to the dimensions of tension, annoyance, negative affect, experience, and tiredness. Regarding the second research question, we were able to prove it by the results of the DDA Questionnaire as the participant noted overall endorsement that the system was a match for their skill due to the adjustment of difficulty of the game was natural and organic, and engaging which indicate to us that FOF across all the experiments using different stimuli was able to adapt the difficulty to the participants.

Another conclusion that can be observed is that the system's usability was considered positive, ranking above the 68th percentile in the system usability scale questionnaire. On the other hand, in terms of the VR experience, the participants were noted to feel the environment was not realistic but engaging and did not cause any severe adverse effects while experiencing it.

7.2 Limitations and Future Work

Although the positive results from this thesis suggest a positive experience and an excellent adaptation to difficulty, the study was plagued by time constraints, making it challenging to test various DDA algorithms, Sensors and games with the FOF. One more aspect limited by the time constraints would be the evaluation of the system usability of the FOF.

If we continue the project, these limitations will be address by implementing additional algorithms for DDA, which provide algorithms to produce a better suit for any situation. Also, the system should be implemented to be compatible with other data transmissions. Furthermore, a more comprehensive range of games and sensors shall be used. This approach will enables a better understanding of the system's adaptability in a myriad of scenarios.

Lastly, a focus usability test should be used to evaluate how intuitive the system is for unfamiliar people. This test should involve approximately 10 to 15 participants tasked with implementing FOF in a game. Data will be collected through questionnaires focusing on the major errors and difficulties encountered. The feedback will be valuable to guide for necessary fixes and improvements to ensure a more user-friendly experience.

7.3 Final Remarks

Optimizing user engagement and learning outcomes is crucial in the dynamic landscape of serious games. DDA stands as a vital tool for tailoring the gaming experience, but the absence of a standardized, platform-agnostic framework limits its widespread implementation. With this thesis, we take a step closer to solving this issue and helping patients perform a more joyful and engaging rehabilitation.

References

[1] T. Huber, S. Mertes, S. Rangelova, S. Flutura, and E. Andre, “Dynamic Difficulty Adjustment in Virtual Reality Exergames through Experience-driven Procedural Content Generation,” in *2021 IEEE Symposium Series on Computational Intelligence (SSCI)*. Orlando, FL, USA: IEEE, Dec. 2021, pp. 1–8. [Online]. Available: https://ieeexplore.ieee.org/document/9660086/

[2] G. Cyrino, J. Tannús, E. Lamounier, A. Cardoso, and A. Soares, “HarpyGame: A Customizable Serious Game for Upper Limb Rehabilitation after Stroke,” in *Anais Estendidos do Simpósio de Realidade Virtual e Aumentada (SVR)*. SBC, Nov. 2020, pp. 67–68, iSSN: 0000-0000. [Online]. Available: https://sol.sbc.org.br/index.php/svr_estendido/article/view/12960

[3] M. Zohaib, “Dynamic Difficulty Adjustment (DDA) in Computer Games: A Review,” *Advances in Human-Computer Interaction*, vol. 2018, pp. 1–12, Nov. 2018. [Online]. Available: https://www.hindawi.com/journals/ahci/2018/5681652/

[4] G. K. Sepulveda, F. Besoain, and N. A. Barriga, “Exploring Dynamic Difficulty Adjustment in Videogames,” in *2019 IEEE CHILEAN Conference on Electrical, Electronics Engineering, Information and Communication Technologies (CHILECON)*, Nov. 2019, pp. 1–6.

[5] P. Jawinski, H. Kirsten, C. Sander, J. Spada, C. Ulke, J. Huang, R. Burkhardt, M. Scholz, T. Hensch, and U. Hegerl, “Human brain arousal in the resting state: a genome-wide association study,” *Molecular Psychiatry*, vol. 24, no. 11, pp. 1599–1609, Nov. 2019. [Online]. Available: https://doi.org/10.1038/s41380-018-0052-2

[6] D. Melhart, A. Liapis, and G. N. Yannakakis, “The arousal video game annotation (again) dataset,” *IEEE Transactions on Affective Computing*, vol. 13, no. 4, pp. 2171–2184, 2022.

[7] R. Lima, D. F. de Noronha Osório, and H. Gamboa, “Heart rate variability and electrodermal activity in mental stress aloud: Predicting the outcome.” in *Biosignals*, 2019, pp. 42–51.

[8] D. Dziedzic and W. Włodarczyk, “Approaches to Measuring the Difficulty of Games in Dynamic Difficulty Adjustment Systems,” *International Journal of Human–Computer Interaction*, vol. 34, no. 8, pp. 707–715, Aug. 2018, publisher: Taylor & Francis _eprint: https://doi.org/10.1080/10447318.2018.1461764. [Online]. Available: https://doi.org/10.1080/10447318.2018.1461764

[9] C. Liu, P. Agrawal, N. Sarkar, and S. Chen, “Dynamic Difficulty Adjustment in Computer Games Through Real-Time Anxiety-Based Affective Feedback,” *International Journal of Human–Computer Interaction*, vol. 25, no. 6, pp. 506–529, Aug. 2009, publisher: Taylor & Francis _eprint: https://doi.org/10.1080/10447310902963944. [Online]. Available: https://doi.org/10.1080/10447310902963944

[10] A. Stein, Y. Yotam, R. Puzis, G. Shani, and M. Taieb-Maimon, “EEG-triggered dynamic difficulty adjustment for multiplayer games,” *Entertainment Computing*, vol. 25, pp. 14–25, Mar. 2018. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1875952117301325

[11] E. Lach, “A Quick Method for Dynamic Difficulty Adjustment of a Computer Player in Computer Games,” in *Artificial Intelligence and Soft Computing*, ser. Lecture Notes in Computer Science, L. Rutkowski, M. Korytkowski, R. Scherer, R. Tadeusiewicz, L. A. Zadeh, and J. M. Zurada, Eds. Cham: Springer International Publishing, 2015, pp. 669–678.

[12] F. Khan, B. Amatya, M. P. Galea, R. Gonzenbach, and J. Kesselring, “Neurorehabilitation: applied neuroplasticity,” *Journal of Neurology*, vol. 264, no. 3, pp. 603–615, Mar. 2017. [Online]. Available: https://doi.org/10.1007/s00415-016-8307-9

[13] A. L. Faria, M. S. Pinho, and S. Bermúdez i Badia, “A comparison of two personalization and adaptive cognitive rehabilitation approaches: a randomized controlled trial with chronic stroke patients,” *Journal of NeuroEngineering and Rehabilitation*, vol. 17, no. 1, p. 78, Jun. 2020. [Online]. Available: https://doi.org/10.1186/s12984-020-00691-5

[14] T. Massetti, T. D. da Silva, T. B. Crocetta, R. Guarnieri, B. L. de Freitas, P. Bianchi Lopes, S. Watson, J. Tonks, and C. B. de Mello Monteiro, “The Clinical Utility of Virtual Reality in Neurorehabilitation: A Systematic Review,” *Journal of Central Nervous System Disease*, vol. 10, p. 117957351881354, Jan. 2018. [Online]. Available: http://journals.sagepub.com/doi/10.1177/1179573518813541

[15] M. Iosa, G. Morone, A. Cherubini, and S. Paolucci, “The Three Laws of Neurorobotics: A Review on What Neurorehabilitation Robots Should Do for Patients and Clinicians,” *Journal of Medical and Biological Engineering*, vol. 36, no. 1, pp. 1–11, Feb. 2016. [Online]. Available: https://doi.org/10.1007/s40846-016-0115-2

[16] J. Tosto-Mancuso, L. Tabacof, J. E. Herrera, E. Breyman, S. Dewil, M. Cortes, L. Correa-esnard, C. P. Kellner, N. Dangayach, and D. Putrino, “Gamified Neurorehabilitation Strategies for Post-stroke Motor Recovery: Challenges and Advantages,” *Current Neurology and Neuroscience Reports*, vol. 22, no. 3, pp. 183–195, Mar. 2022. [Online]. Available: https://doi.org/10.1007/s11910-022-01181-y

[17] S. Bermúdez i Badia, J. Avelino, A. Bernardino, M. S. Cameirão, J. E. Muñoz, H. Cardoso, A. Gonçalves, T. Paulino, R. Ribeiro, H. Simão *et al.*, “Development and validation of a mixed reality exergaming platform for fitness training of older adults,” in *Everyday Virtual and Augmented Reality*. Springer, 2023, pp. 119–145.

[18] T. Paulino, A. L. Faria, and S. Bermúdez i Badia, “Reh@City v2.0: a comprehensive virtual reality cognitive training system based on personalized and adaptive simulations of activities of daily living,” in *2019 5th Experiment International Conference (exp.at’19)*, Jun. 2019, pp. 292–297. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/8876539

[19] A. Vourvopoulos, A. L. Faria, M. S. Cameirão, and S. Bermúdez i Badia, “RehabNet: A distributed architecture for motor and cognitive neuro-rehabilitation,” pp. 454–459, Oct. 2013. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/6720719

[20] G. G. Fluet, D. Roy, R. Llorens, S. Bermúdez i Badia, and J. E. Deutsch, “Basis and Clinical Evidence of Virtual Reality-Based Rehabilitation of Sensorimotor Impairments After Stroke,” in *Neurorehabilitation Technology*, D. J. Reinkensmeyer, L. Marchal-Crespo, and V. Dietz, Eds. Cham: Springer International Publishing, 2022, pp. 429–466. [Online]. Available: https://link.springer.com/10.1007/978-3-031-08995-4_20

[21] D. Cano Porras, H. Sharon, R. Inzelberg, Y. Ziv-Ner, G. Zeilig, and M. Plotnik, "Advanced virtual reality-based rehabilitation of balance and gait in clinical practice," *Therapeutic Advances in Chronic Disease*, vol. 10, p. 2040622319868379, Jan. 2019, publisher: SAGE Publications. [Online]. Available: https://doi.org/10.1177/2040622319868379

[22] M. Lee, D. Suh, J. Son, J. Kim, S.-D. Eun, and B. Yoon, "Patient perspectives on virtual reality-based rehabilitation after knee surgery: Importance of level of difficulty," *Journal of Rehabilitation Research and Development*, vol. 53, no. 2, pp. 239–252, 2016. [Online]. Available: http://www.rehab.research.va.gov/jour/2016/532/pdf/JRRD-2014-07-0164.pdf

[23] N. Aderinto, G. Olatunji, M. O. Abdulbasit, M. Edun, G. Aboderin, and E. Egbunu, "Exploring the efficacy of virtual reality-based rehabilitation in stroke: a narrative review of current evidence," *Annals of Medicine*, vol. 55, no. 2, p. 2285907, Dec. 2023. [Online]. Available: https://www.tandfonline.com/doi/full/10.1080/07853890.2023.2285907

[24] S.-H. Kwon, J. K. Park, and Y. H. Koh, "A systematic review and meta-analysis on the effect of virtual reality-based rehabilitation for people with Parkinson's disease," *Journal of NeuroEngineering and Rehabilitation*, vol. 20, no. 1, p. 94, Jul. 2023. [Online]. Available: https://doi.org/10.1186/s12984-023-01219-3

[25] B. Rubio Ballester, S. Bermúdez i Badia, and P. F. Verschure, "The effect of social gaming in virtual reality based rehabilitation of stroke patients," in *2011 International Conference on Virtual Rehabilitation*, 2011, pp. 1–2.

[26] M. S. Cameirão, S. B. i. Badia, E. D. Oller, and P. F. Verschure, "Neurorehabilitation using the virtual reality based Rehabilitation Gaming System: methodology, design, psychometrics, usability and validation," *Journal of NeuroEngineering and Rehabilitation*, vol. 7, no. 1, p. 48, Sep. 2010. [Online]. Available: https://doi.org/10.1186/1743-0003-7-48

[27] "Fig. 1. The Rehabilitation Gaming System setup. The user wears color..." [Online]. Available: https://www.researchgate.net/figure/The-Rehabilitation-Gaming-System-setup-The-user-wears-color-patches-that-are-tracked-by_fig2_4283552

[28] M. Pezzera and N. A. Borghese, "Dynamic difficulty adjustment in exer-games for rehabilitation: a mixed approach," in *2020 IEEE 8th International Conference on Serious Games and Applications for Health (SeGAH)*. Vancouver, BC, Canada: IEEE, Aug. 2020, pp. 1–7. [Online]. Available: https://ieeexplore.ieee.org/document/9201871/

[29] K. d. O. Andrade, T. B. Pasqual, G. A. P. Caurin, and M. K. Crocomo, "Dynamic difficulty adjustment with Evolutionary Algorithm in games for rehabilitation robotics," in *2016 IEEE International Conference on Serious Games and Applications for Health (SeGAH)*. Orlando, FL, USA: IEEE, May 2016, pp. 1–8. [Online]. Available: http://ieeexplore.ieee.org/document/7586277/

[30] C. H. Souza, S. S. De Oliveira, L. O. Berretta, and S. T. De Carvalho, "DDA-MAPEKit: A Framework for Dynamic Difficulty Adjustment Based on MAPE-K Loop," *ACM International Conference Proceeding Series*, pp. 1–10, 2023.

[31] J. Hagelback and S. J. Johansson, "Measuring player experience on runtime dynamic difficulty scaling in an RTS game," in *2009 IEEE Symposium on Computational Intelligence and Games*, Sep. 2009, pp. 46–52, iSSN: 2325-4289.

[32] C. Pedersen, J. Togelius, and G. N. Yannakakis, "Modeling player experience in Super Mario Bros," in *2009 IEEE Symposium on Computational Intelligence and Games*, Sep. 2009, pp. 132–139, iSSN: 2325-4289.

[33] X. Li, S. He, Y. Dong, Q. Liu, X. Liu, Y. Fu, Z. Shi, and W. Huang, "To create DDA by the approach of ANN from UCT-created data," in *2010 International Conference on Computer Application and System Modeling (ICCASM 2010)*, vol. 8, Oct. 2010, pp. V8–475–V8–478, iSSN: 2161-9077.

[34] P. Spronck, M. Ponsen, I. Sprinkhuizen-Kuyper, and E. Postma, "Adaptive game AI with dynamic scripting," *Machine Learning*, vol. 63, no. 3, pp. 217–248, Jun. 2006. [Online]. Available: https://doi.org/10.1007/s10994-006-6205-6

[35] A. Ebrahimi and M.-R. Akbarzadeh-T, "Dynamic difficulty adjustment in games by using an interactive self-organizing architecture," in *2014 Iranian Conference on Intelligent Systems (ICIS)*, Feb. 2014, pp. 1–6.

[36] "Biofeedback: Types, Purpose, and Risks," Jul. 2012. [Online]. Available: https://www.healthline.com/health/biofeedback

[37] S. Shapoval, B. García Zapirain, A. Mendez Zorrilla, and I. Mugueta-Aguinaga, "Biofeedback Applied to Interactive Serious Games to Monitor Frailty in an Elderly Population," *Applied Sciences*, vol. 11, no. 8, p. 3502, Apr. 2021. [Online]. Available: https://www.mdpi.com/2076-3417/11/8/3502

[38] V. Gohel and N. Mehendale, "Review on electromyography signal acquisition and processing," *Biophysical Reviews*, vol. 12, no. 6, pp. 1361–1367, Dec. 2020. [Online]. Available: https://doi.org/10.1007/s12551-020-00770-w

[39] O. M. Giggins, U. M. Persson, and B. Caulfield, "Biofeedback in rehabilitation," *Journal of NeuroEngineering and Rehabilitation*, vol. 10, no. 1, p. 60, Jun. 2013. [Online]. Available: https://doi.org/10.1186/1743-0003-10-60

[40] "Biofeedback - an overview | ScienceDirect Topics." [Online]. Available: https://www.sciencedirect.com/topics/neuroscience/biofeedback

[41] M. Sundaram, R. Sadler, G. Young, and N. Pillay, "EEG in Epilepsy: Current Perspectives," *Canadian Journal of Neurological Sciences / Journal Canadien des Sciences Neurologiques*, vol. 26, no. 4, pp. 255–262, Nov. 1999. [Online]. Available: https://www.cambridge.org/core/product/identifier/S0317167100000342/type/journal_article

[42] P. Antonenko, F. Paas, R. Grabner, and T. van Gog, "Using Electroencephalography to Measure Cognitive Load," *Educational Psychology Review*, vol. 22, no. 4, pp. 425–438, Dec. 2010. [Online]. Available: https://doi.org/10.1007/s10648-010-9130-y

[43] M. Dias da Silva, Gonçalves, D. Branco, and M. Postma, "Revisiting consciousness: Distinguishing between states of conscious focused attention and mind wandering with EEG," *Consciousness and Cognition*, vol. 101, p. 103332, May 2022. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S1053810022000642

[44] R. H. Stevens, T. Galloway, and C. Berka, "EEG-Related Changes in Cognitive Workload, Engagement and Distraction as Students Acquire Problem Solving Skills," in *User Modeling 2007*, C. Conati, K. McCoy, and G. Paliouras, Eds. Berlin,

Heidelberg: Springer Berlin Heidelberg, 2007, vol. 4511, pp. 187–196, iSSN: 0302-9743, 1611-3349 Series Title: Lecture Notes in Computer Science. [Online]. Available: http://link.springer.com/10.1007/978-3-540-73078-1_22

[45] G. Fortino and V. Giampà, "PPG-based methods for non invasive and continuous blood pressure measurement: an overview and development issues in body sensor networks," in *2010 IEEE International Workshop on Medical Measurements and Applications*, Apr. 2010, pp. 10–13.

[46] R. Hunicke, "The case for dynamic difficulty adjustment in games," in *Proceedings of the 2005 ACM SIGCHI International Conference on Advances in computer entertainment technology*, ser. ACE '05. New York, NY, USA: Association for Computing Machinery, Jun. 2005, pp. 429–433. [Online]. Available: https://doi.org/10.1145/1178477.1178573

[47] C. Liu, P. Agrawal, N. Sarkar, and S. Chen, "Dynamic Difficulty Adjustment in Computer Games Through Real-Time Anxiety-Based Affective Feedback," *International Journal of Human–Computer Interaction*, vol. 25, no. 6, pp. 506–529, Aug. 2009, publisher: Taylor & Francis _eprint: https://doi.org/10.1080/10447310902963944. [Online]. Available: https://doi.org/10.1080/10447310902963944

[48] M. Jennings-Teats, G. Smith, and N. Wardrip-Fruin, "Polymorph: dynamic difficulty adjustment through level generation," in *Proceedings of the 2010 Workshop on Procedural Content Generation in Games*, ser. PCGames '10. New York, NY, USA: Association for Computing Machinery, Jun. 2010, pp. 1–4. [Online]. Available: https://doi.org/10.1145/1814256.1814267

[49] Q. Wang, Q. Zhang, W. Sun, C. Boulay, K. Kim, and R. L. Barmaki, "A scoping review of the use of lab streaming layer framework in virtual and augmented reality research," *Virtual Reality*, vol. 27, no. 3, pp. 2195–2210, Sep. 2023. [Online]. Available: https://doi.org/10.1007/s10055-023-00799-8

[50] K. C, "LabStreamingLayer,https://github.com/sccn/labstreaminglayer," 2023. [Online]. Available: https://github.com/sccn/labstreaminglayer

[51] "OpenSignals." [Online]. Available: https://www.pluxbiosignals.com/pages/opensignals

[52] M. Gnacek, J. Broulidakis, I. Mavridou, M. Fatoorechi, E. Seiss, T. Kostoulas, E. Balaguer-Ballester, I. Kiprijanovska, C. Rosten, and C. Nduka, "emteqPRO—Fully Integrated Biometric Sensing Array for Non-Invasive Biomedical Research in Virtual Reality," *Frontiers in Virtual Reality*, vol. 3, 2022. [Online]. Available: https://www.frontiersin.org/articles/10.3389/frvir.2022.781218

[53] W. A. IJsselsteijn, Y. A. De Kort, and K. Poels, "The game experience questionnaire," 2013.

[54] J. Brooke, "Sus: a "quick and dirty'usability," vol. 189, no. 3, pp. 189–194, publisher: Taylor & Francis.

[55] J. Vasconcelos-Raposo, M. Melo, C. Teixeira, L. Cabral, and M. Bessa, "Adaptation and validation of the ITC - Sense of Presence Inventory for the Portuguese language," *International Journal of Human-Computer Studies*, vol. 125, pp. 1–6, May 2019. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S1071581918306840

[56] B. G. Witmer and M. J. Singer, "Measuring presence in virtual environments: A presence questionnaire," *Presence*, vol. 7, no. 3, pp. 225–240, 1998.

[57] "Polar H10 Heart Rate Sensor M-XXL Unisex." [Online]. Available: https://21run.com/eu/polar-h10-heart-rate-sensor-m-xxl-unisex-920106245

[58] T. Rose, C. S. Nam, and K. B. Chen, "Immersion of virtual reality for rehabilitation - Review," *Applied Ergonomics*, vol. 69, pp. 153–161, May 2018. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S0003687018300176

8 Annexes

8.1 Annex A

NeuroRehabLab

Manual For the use of:

Flow Optimizer Framework

Pyflow:

Installation:

---

1. Clone or download repository [(https://github.com/NeuroRehabilitation/PyFlow).](https://github.com/NeuroRehabilitation/PyFlow)
2. Install all the necessary libraries by running the starterUI.bat or by installing all the requirements in the file requirement.txt.

Overview:

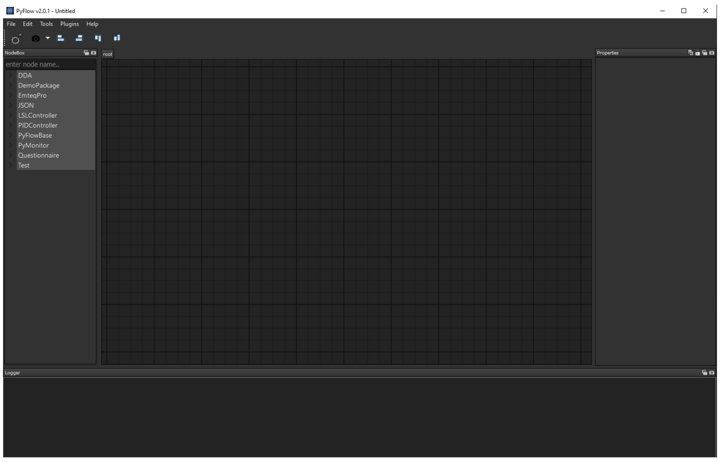
---

1. Official website: https://wonderworks-software.github.io/PyFlow/
2. Documentation: https://pyflow.readthedocs.io/en/latest/

Implementation:

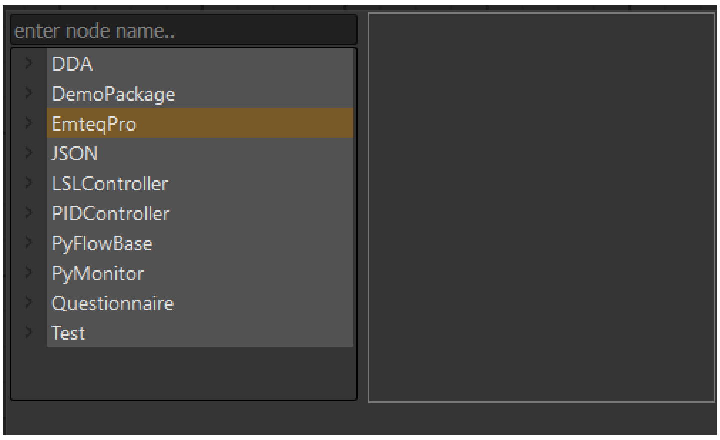
Open Pyflow:

1. To run the program in standalone mode, run pyflow.py. It can be found in the root folder of the repository.



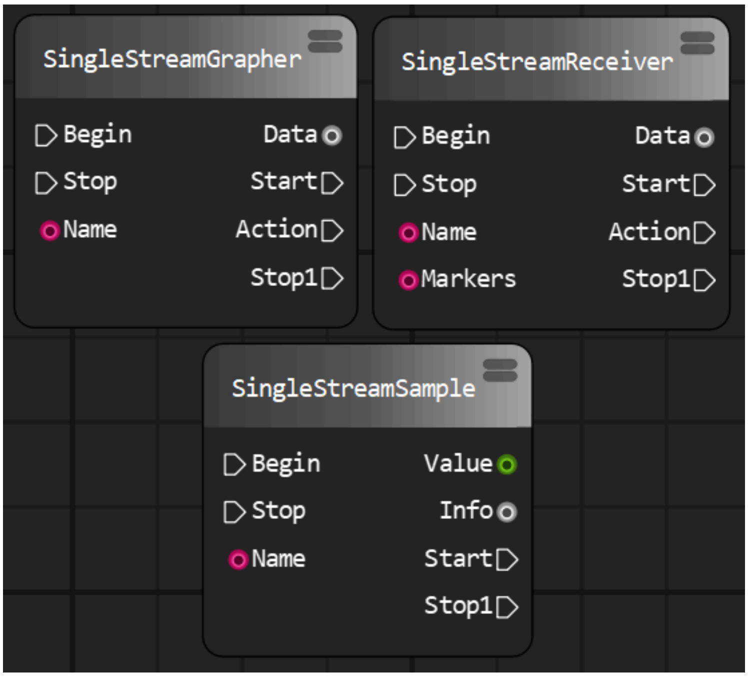
Add Node:

1. Press the right mouse button.
2. Choose which node to add.



Pyflow:

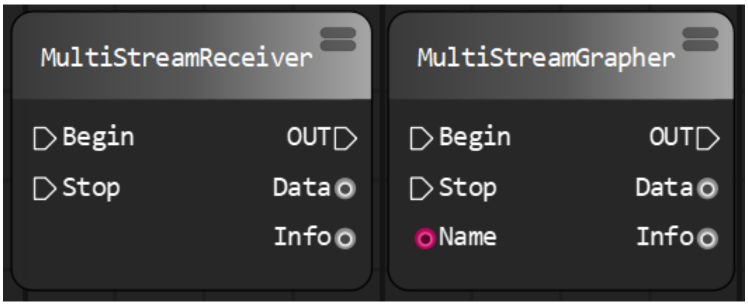
LSL Input Nodes:



1. **SingleStreamReceiver:** Gets data from one stream and sends the latest one-second data to connected nodes.

2. **SingleStreamGrapher:** Gets data from one stream and sends the latest one-second data to connected nodes. Also displays the data graphically using PyMonitor.

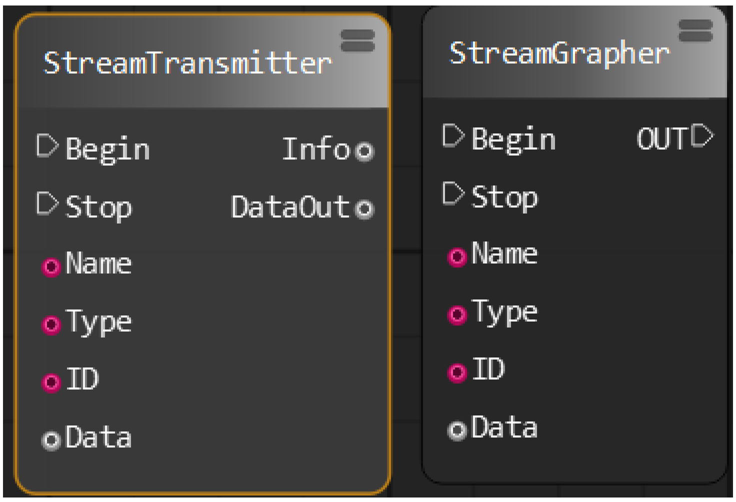
3. **SingleStreamSample:** Gets data from one stream and sends the latest one-second data to connected nodes. Also displays the data graphically using PyMonitor.



4. **MultiStreamReceiver:** Gets data from all streams and sends the latest one-second data from each stream to connected nodes.

5. **MultiStreamGrapher:** Gets data from all streams and sends the latest one-second data from each stream to connected nodes. Also displays the data graphically using PyMonitor.

LSL Output Nodes:



1. **StreamTransmitter**: Sets up an LSL data stream to transmit received information from the node.

2. **StreamGrapher:** Sets up an LSL data stream to transmit received information from the node. Additionally, it displays the sent data in graphs using PyMonitor.

Pyflow:

PID Controller Nodes:

1. **PID Node 1:**This node utilizes the PID algorithm to calculate, without prior data, the optimal changes in difficulty for the player during their playthrough. It does so based on a SetPoint, which represents the objective that the player should achieve, the player's performance, and the constants KP, KI, and KD. These constants aid in reaching the correct difficulty level and prevent overshooting.

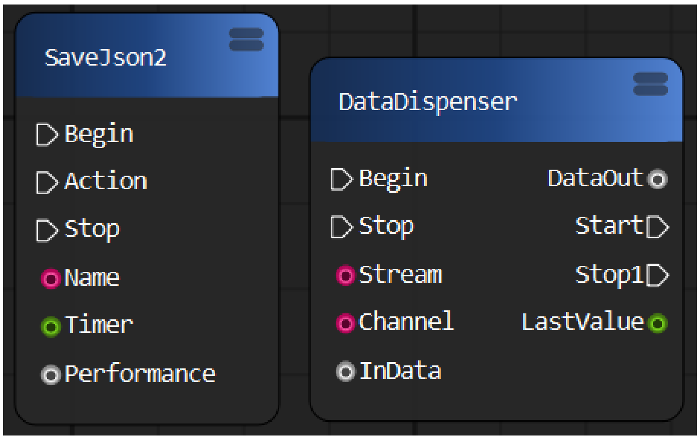
2. **PID Node 2:** This node utilizes the PID algorithm to calculate, without prior data, the optimal changes in difficulty for the player during their playthrough. It does so based on a SetPoint, which represents the objective that the player should achieve, the player's performance, and the constants KP, KI, and KD. These constants aid in reaching the correct difficulty level and prevent overshooting.

Pyflow:

Misc Nodes:

1. **SaveJSON:** Saves data into a JSON file.

2. **Data Dispenser:** A node designed to filter dictionary data and allows retrieval of all data as an array or just the last received value, providing greater flexibility in node usage.



Create New Node:

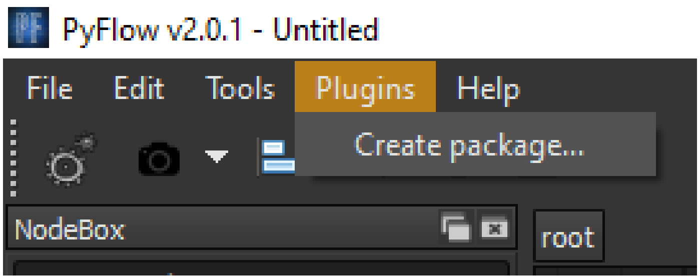
1. Select the Create Package option in the plugins menu.

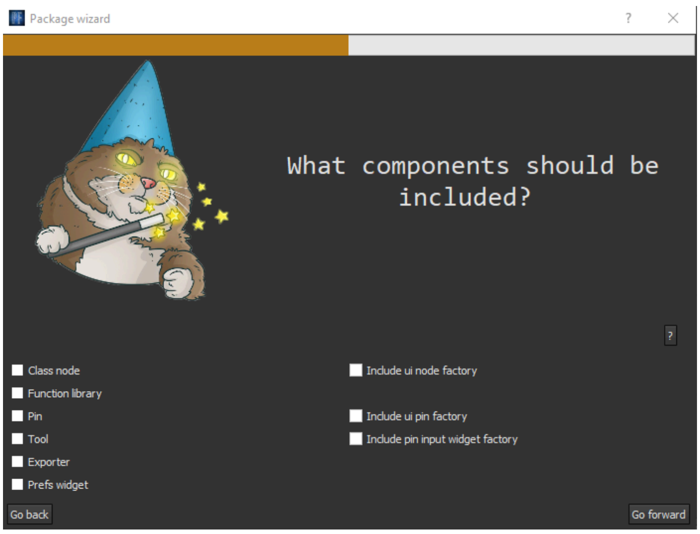
2. Press the forward button.

3. Choose a name for the package.

4. Choose the components needed for the package.

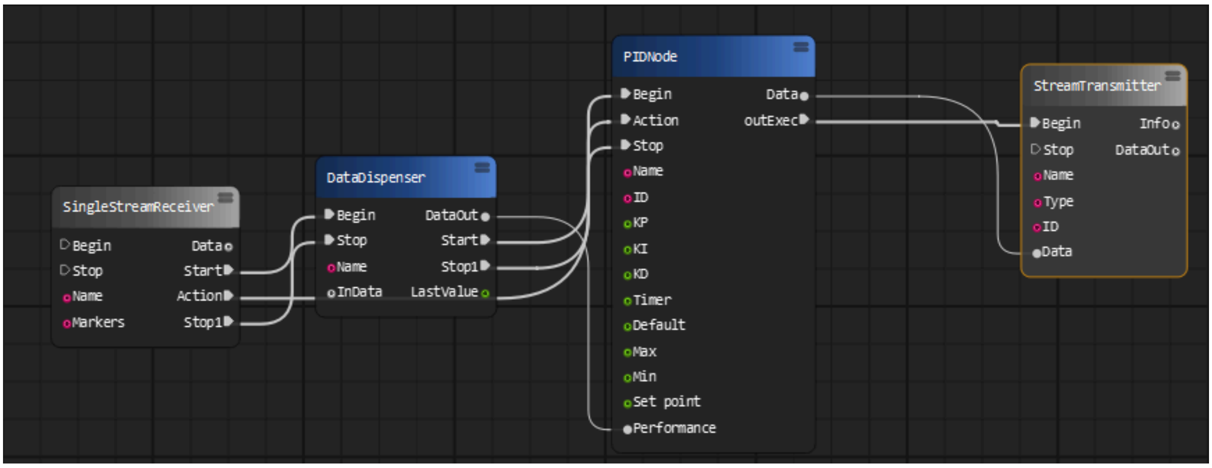
5. Choose the directory to save the package.





Pyflow:

FOF Example:

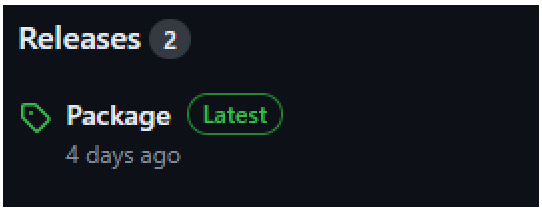


DDA Framework:

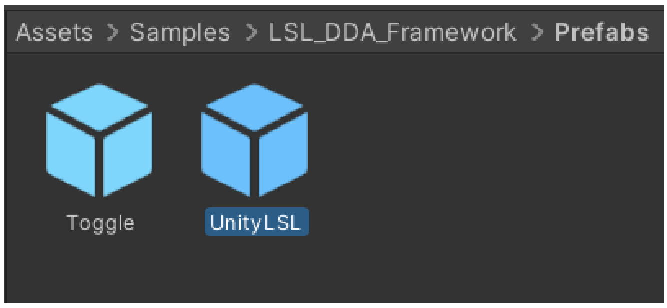
Installation:

---

1. Download package(https://github.com/Pedro3Lobo/Unity_LSL_Framework).



2. Import your package into your Unity project.

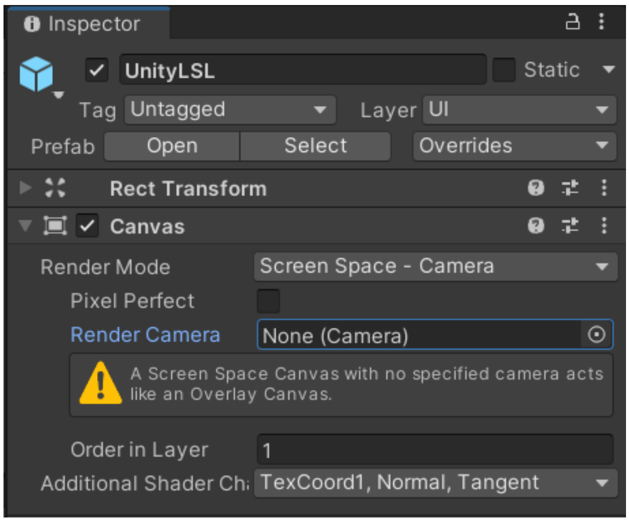


3. Drag and drop the provided UnityLSL prefab into your scene.

Implementation:

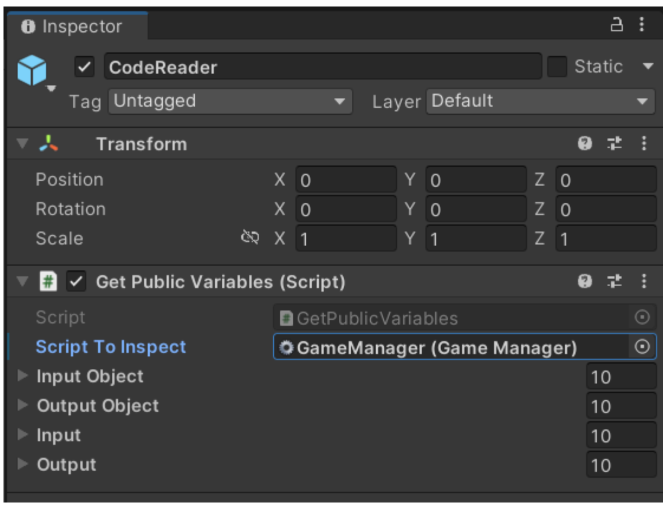
Add Control Panel to Camera:

1. In the UnityLSL object in the scene, select the canvas component and add the Main Camera to the Render Camera.



Add Variables Source Script:

1. Select in scene the child of Unity LSL and Manager, CodeReader Object and Drag and Drop a script in the Script To Inspect input.

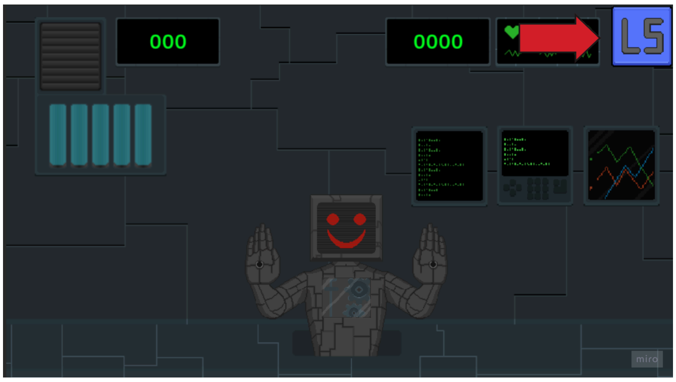


DDA Framework:

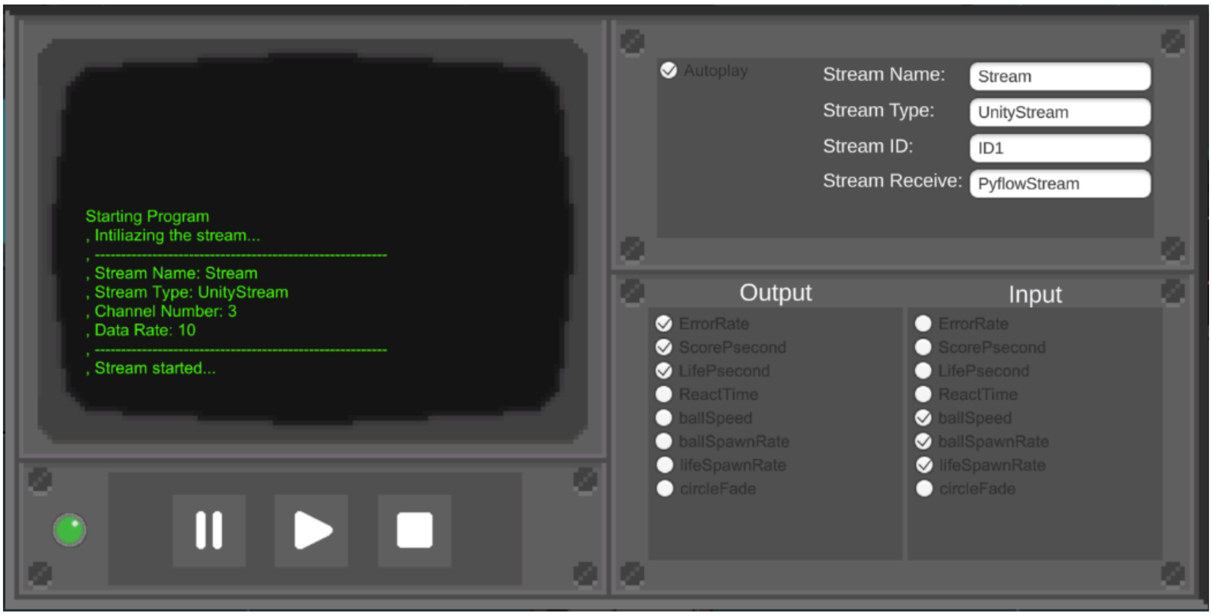
Access the Control Panel:

1. Play the Unity Scene.

2. Press the button in the upper left side of the screen and select.

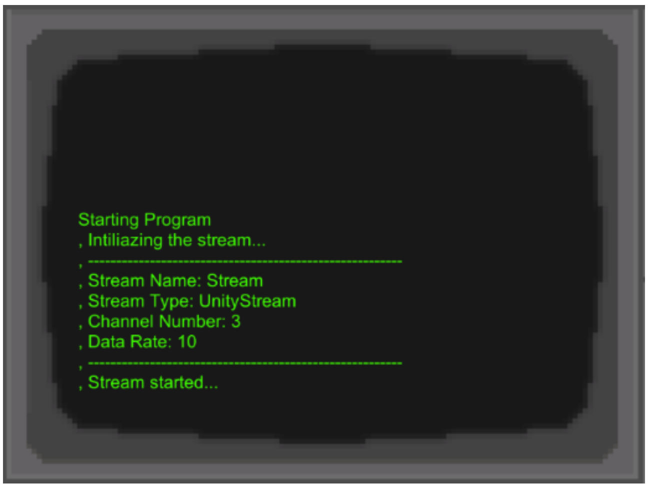


Control Panel:



Console Display:

This section provides details about the created stream, including its name, type, number of channels, and data rate.



8.2 Annex B

TQ-GamingBackground

* Indica uma pergunta obrigatória

Participant Consent Form

Thank you for considering participating in our study! Before you proceed, we kindly ask for your consent to use your data. Please read the following information carefully:

1. **Confidentiality:** All the information collected during this study will be treated with strict confidentiality. Your personal information will be anonymized and stored securely.
2. **Voluntary Participation:** Participation in this study is entirely voluntary. You have the right to withdraw at any time without providing a reason, and this will not have any negative consequences for you.
3. **Purpose of the Study:** The purpose of this study is to assess how individuals respond to varying levels of difficulty. Your valuable participation will greatly contribute to advancing our understanding of game limits and help us better adapt the difficulty levels to enhance the gaming experience.
4. **Data Protection:** We assure you that your data will be handled in compliance with data protection laws and regulations. Your information will only be used for the purposes of this study and will not be shared with any third parties.
5. **Contact Information:** If you have any questions or concerns about the study, you can contact our research team at pedrolobowork@gmail.com.

By clicking the "I Consent" button below, you confirm that you have read and understood the information provided and voluntarily consent to participate in this study.

1. Please enter your number below: *

____

2. *

Marcar apenas uma oval.

- ( ) I Consent
- ( ) I Decline

Games Background

In this section of the form, we would like to gather information about your background in the world of video games and understand your gaming preferences in more detail.

3. Gender? *

Marcar apenas uma oval.

- Male
- Female
- Outra: ____________

4. Age Group? *

Marcar apenas uma oval.

- 18-30
- 31-39
- 40-49
- 50 or Over

5. How much experience do you have playing video games? *

Marcar apenas uma oval.

	1	2	3	4	5	
Non						Alot

6. What difficulty level do you usually choose when playing games? *

Marcar apenas uma oval.

- Easy
- Normal
- Hard

7. How often do you play video games? (daily, weekly, monthly) *

Marcar apenas uma oval.

- Daily
- Weekly
- Monthly
- Never *Avançar para a pergunta 9*

8. On average, how many hours do you spend gaming per session? *

Marcar apenas uma oval.

- Less then 1 hour
- 1-3 hours
- 3-6 hours
- 6-9 hours
- I don't play games

Games Background(Cont...)

9. What genre or types of games do you typically prefer to play? *

Marcar apenas uma oval.

- Action
- Adventure
- Role-playing
- Strategy
- Sports
- Puzzle
- Racing
- Simulation
- Fighting
- Outra: ____________________

10. Are there any genres you avoid? *

Marcar apenas uma oval.

- Action
- Adventure
- Role-playing
- Strategy
- Sports
- Puzzle
- Racing
- Simulation
- Fighting
- Outra: ____________

11. Do you ever attempted to play games at the highest difficulty level when given the opportunity? *

Marcar apenas uma oval.

	1	2	3	4	5	
Nev	◯	◯	◯	◯	◯	Always

12. Do you ever attempted to play games at the normal difficulty level when given the opportunity? *

Marcar apenas uma oval.

	1	2	3	4	5	
Nev	◯	◯	◯	◯	◯	Always

13. Do you ever attempted to play games at the lowest difficulty level when given the opportunity? *

Marcar apenas uma oval.

	1	2	3	4	5	
Nev	◯	◯	◯	◯	◯	Always

8.3 Annex C

GE_Core

Please indicate how you felt **while** playing the game for each of the items, on the following scale:

- 0-not at all
- 1-slightly
- 2-moderately
- 3-fairly
- 4-extremely

* Indica uma pergunta obrigatória

1. Please enter your number below:

____

2. 1.I felt content *

Marcar apenas uma oval.

	0	1	2	3	4	
not a	◯	◯	◯	◯	◯	extremely

3. 2.I felt skilful *

Marcar apenas uma oval.

	0	1	2	3	4	
not a	◯	◯	◯	◯	◯	extremely

4. 3.I was interested in the game's story *

Marcar apenas uma oval.

	0	1	2	3	4	
not ;	◯	◯	◯	◯	◯	extremely

5. 4.I thought it was fun *

Marcar apenas uma oval.

	0	1	2	3	4	
not ;	◯	◯	◯	◯	◯	extremely

6. 5.I was fully occupied with the game *

Marcar apenas uma oval.

	0	1	2	3	4	
not ;	◯	◯	◯	◯	◯	extremely

7. 6.I felt happy *

Marcar apenas uma oval.

	0	1	2	3	4	
not ;	◯	◯	◯	◯	◯	extremely

8. 7.It gave me a bad mood *

Marcar apenas uma oval.

	0	1	2	3	4	
not a	◯	◯	◯	◯	◯	extremely

9. 8.I thought about other things *

Marcar apenas uma oval.

	0	1	2	3	4	
not a	◯	◯	◯	◯	◯	extremely

10. 9.I found it tiresome *

Marcar apenas uma oval.

	0	1	2	3	4	
not a	◯	◯	◯	◯	◯	extremely

11. 10.I felt competent *

Marcar apenas uma oval.

	0	1	2	3	4	
not a	◯	◯	◯	◯	◯	extremely

12. 11.I thought it was hard *

Marcar apenas uma oval.

	0	1	2	3	4	
not a	◯	◯	◯	◯	◯	extremely

13. 12.It was aesthetically pleasing *

Marcar apenas uma oval.

	0	1	2	3	4	
not a	◯	◯	◯	◯	◯	extremely

14. 13.I forgot everything around me *

Marcar apenas uma oval.

	0	1	2	3	4	
not a	◯	◯	◯	◯	◯	extremely

15. 14.I felt good *

Marcar apenas uma oval.

	0	1	2	3	4	
not a	◯	◯	◯	◯	◯	extremely

16. 15.I was good at it *

Marcar apenas uma oval.

	0	1	2	3	4	
not a	◯	◯	◯	◯	◯	extremely

17. 16.I felt bored *

Marcar apenas uma oval.

	0	1	2	3	4	
not a	◯	◯	◯	◯	◯	extremely

18. 17.I felt successful *

Marcar apenas uma oval.

	0	1	2	3	4	
not a	◯	◯	◯	◯	◯	extremely

19. 18.I felt imaginative *

Marcar apenas uma oval.

	0	1	2	3	4	
not a	◯	◯	◯	◯	◯	extremely

20. 19.I felt that I could explore things *

Marcar apenas uma oval.

	0	1	2	3	4	
not ;	◯	◯	◯	◯	◯	extremely

21. 20.I enjoyed it *

Marcar apenas uma oval.

	0	1	2	3	4	
not ;	◯	◯	◯	◯	◯	extremely

22. 21.I was fast at reaching the game's targets *

Marcar apenas uma oval.

	0	1	2	3	4	
not ;	◯	◯	◯	◯	◯	extremely

23. 22.I felt annoyed *

Marcar apenas uma oval.

	0	1	2	3	4	
not ;	◯	◯	◯	◯	◯	extremely

24. 23.I felt pressured *

Marcar apenas uma oval.

	0	1	2	3	4	
not a	◯	◯	◯	◯	◯	extremely

25. 24.I felt irritable *

Marcar apenas uma oval.

	0	1	2	3	4	
not a	◯	◯	◯	◯	◯	extremely

26. 25.I lost track of time *

Marcar apenas uma oval.

	0	1	2	3	4	
not a	◯	◯	◯	◯	◯	extremely

27. 26.I felt challenged *

Marcar apenas uma oval.

	0	1	2	3	4	
not a	◯	◯	◯	◯	◯	extremely

28. 27.I found it impressive *

Marcar apenas uma oval.

	0	1	2	3	4	
not a	◯	◯	◯	◯	◯	extremely

29. 28.I was deeply concentrated in the game *

Marcar apenas uma oval.

	0	1	2	3	4	
not a	◯	◯	◯	◯	◯	extremely

30. 30.I felt frustrated *

Marcar apenas uma oval.

	0	1	2	3	4	
not a	◯	◯	◯	◯	◯	extremely

31. 31.It felt like a rich experience *

Marcar apenas uma oval.

	0	1	2	3	4	
not a	◯	◯	◯	◯	◯	extremely

32. 32.I lost connection with the outside world *

Marcar apenas uma oval.

	0	1	2	3	4	
not a	◯	◯	◯	◯	◯	extremely

33. 33.I had to put a lot of effort into it *

Marcar apenas uma oval.

	0	1	2	3	4	
not a	◯	◯	◯	◯	◯	extremely

8.4 Annex D

GE_PG

Please indicate how you felt **after** you finished playing the game for each of the items, on the following scale:

- 0-not at all
- 1-slightly
- 2-moderately
- 3-fairly
- 4-extremely

* Indica uma pergunta obrigatória

1. Please enter your number below:

____

2. 1.I felt revived *

Marcar apenas uma oval.

	0	1	2	3	4	
not a	◯	◯	◯	◯	◯	extremely

3. 2. I felt bad *

Marcar apenas uma oval.

	0	1	2	3	4	
not a	◯	◯	◯	◯	◯	extremely

4. 3. I found it hard to get back to reality *

Marcar apenas uma oval.

	0	1	2	3	4	
not a	◯	◯	◯	◯	◯	extremely

5. 4. I felt guilty *

Marcar apenas uma oval.

	0	1	2	3	4	
not a	◯	◯	◯	◯	◯	extremely

6. 5. It felt like a victory *

Marcar apenas uma oval.

	0	1	2	3	4	
not a	◯	◯	◯	◯	◯	extremely

7. 6. I found it a waste of time *

Marcar apenas uma oval.

	0	1	2	3	4	
not a	◯	◯	◯	◯	◯	extremely

8. 7. I felt energised *

Marcar apenas uma oval.

	0	1	2	3	4	
not a	◯	◯	◯	◯	◯	extremely

9. 8. I felt satisfied *

Marcar apenas uma oval.

	0	1	2	3	4	
not a	◯	◯	◯	◯	◯	extremely

10. 9. I felt disoriented *

Marcar apenas uma oval.

	0	1	2	3	4	
not a	◯	◯	◯	◯	◯	extremely

11. 10. I felt exhausted *

Marcar apenas uma oval.

	0	1	2	3	4	
not a	◯	◯	◯	◯	◯	extremely

12. 11. I felt that I could have done more useful things *

Marcar apenas uma oval.

	0	1	2	3	4	
not a	◯	◯	◯	◯	◯	extremely

13. 12. I felt powerful *

Marcar apenas uma oval.

	0	1	2	3	4	
not a	◯	◯	◯	◯	◯	extremely

14. 13. I felt weary *

Marcar apenas uma oval.

	0	1	2	3	4	
not a	◯	◯	◯	◯	◯	extremely

15. 14. I felt regret *

Marcar apenas uma oval.

	0	1	2	3	4	
not a	◯	◯	◯	◯	◯	extremely

16. 15. I felt ashamed *

Marcar apenas uma oval.

	0	1	2	3	4	
not a	◯	◯	◯	◯	◯	extremely

17. 16. I felt proud *

Marcar apenas uma oval.

	0	1	2	3	4	
not a	◯	◯	◯	◯	◯	extremely

18. 17. I had a sense that I had returned from a journey *

Marcar apenas uma oval.

	0	1	2	3	4	
not a	◯	◯	◯	◯	◯	extremely

8.5 Annex E

DDA Questionnaire

Please indicate how you felt **while** playing the game for each of the items, on the following scale:

- 0-not at all
- 1-slightly
- 2-moderately
- 3-fairly
- 4-extremely

* Indica uma pergunta obrigatória

1. Please enter your number below: *

2. 1. I felt that the challenge match my skills. *

Marcar apenas uma oval.

	0	1	2	3	4	
not a	◯	◯	◯	◯	◯	extremely

3. 2. I expected a lesser challenge. *

Marcar apenas uma oval.

	0	1	2	3	4	
not a	◯	◯	◯	◯	◯	extremely

4. 3. I expected a greater challenge. *

Marcar apenas uma oval.

	0	1	2	3	4	
not a	◯	◯	◯	◯	◯	extremely

5. 4. I felt like the difficulty changed to quickly. *

Marcar apenas uma oval.

	0	1	2	3	4	
not a	◯	◯	◯	◯	◯	extremely

6. 5. I did notice any difference between the difficulty from beginning to end. *

Marcar apenas uma oval.

	0	1	2	3	4	
not a	◯	◯	◯	◯	◯	extremely

7. 6. I felt like the difficulty changed to slowly. *

Marcar apenas uma oval.

	0	1	2	3	4	
not a	◯	◯	◯	◯	◯	extremely

8. 7. I felt the adjustment of dificulty disrupted the flow of the game. *

Marcar apenas uma oval.

	0	1	2	3	4	
not a	◯	◯	◯	◯	◯	extremely

9. 8. I feel frustrated with the game's difficulty level. *

Marcar apenas uma oval.

	0	1	2	3	4	
not a	◯	◯	◯	◯	◯	extremely

10. 9. I felt that the game adapted to my progress in a way that keeps me engaged. *

Marcar apenas uma oval.

	0	1	2	3	4	
not a	◯	◯	◯	◯	◯	extremely

11. 10. The difficulty adjustments felt organic and natural. *

Marcar apenas uma oval.

	0	1	2	3	4	
not a	◯	◯	◯	◯	◯	extremely

8.6 Annex F

Version Questionnaire

* Indica uma pergunta obrigatória

1. Please enter your number below: *

2. 1.Which game did you enjoy the most? *

   Marcar apenas uma oval.

   - Version 1
   - Version 2
   - Version 3

3. Why? *

4. 2.Which game did you find the most challenging? *

   Marcar apenas uma oval.

   - Version 1
   - Version 2
   - Version 3

5. Why? *

6. 3.Which game would you recommend to others? *

Marcar apenas uma oval.

- Version 1
- Version 2
- Version 3

7. Why? *

8. 4.Which game was the hardest? *

Marcar apenas uma oval.

- Version 1
- Version 2
- Version 3

9. Why? *

10. 5.Which game was the easiest? *

Marcar apenas uma oval.

- Version 1
- Version 2
- Version 3

11. Why? *

12. 6.Which game was more of a match for your skills? *

Marcar apenas uma oval.

- Version 1
- Version 2
- Version 3

13. Why?

____
____
____
____
____

14. 7.Which game was less of a match for your skills? *

Marcar apenas uma oval.

- ( ) Version 1
- ( ) Version 2
- ( ) Version 3

15. Why?

____
____
____
____
____

8.7 Annex G

System Usability Scale

* Indica uma pergunta obrigatória

1. Please enter your number below:

2. 1-I think that I would like to use this system frequently. *

Marcar apenas uma oval.

	1	2	3	4	5	6	7	
Stro	◯	◯	◯	◯	◯	◯	◯	Strongly Agree

3. 2-I found the system unnecessarily complex. *

Marcar apenas uma oval.

	1	2	3	4	5	6	7	
Stro	◯	◯	◯	◯	◯	◯	◯	Strongly Agree

4. 3-I thought the system was easy to use. *

Marcar apenas uma oval.

	1	2	3	4	5	6	7	
Stro	◯	◯	◯	◯	◯	◯	◯	Strongly Agree

5. 4-I think that I would need the support of a technical person to be able to use this system. *

Marcar apenas uma oval.

	1	2	3	4	5	6	7	
Stro	◯	◯	◯	◯	◯	◯	◯	Strongly Agree

6. 5-I found the various functions in this system were well integrated. *

Marcar apenas uma oval.

	1	2	3	4	5	6	7	
Stro	◯	◯	◯	◯	◯	◯	◯	Strongly Agree

7. 6-I believe there was a lot of inconsistency in this system *

Marcar apenas uma oval.

	1	2	3	4	5	6	7	
Stro	◯	◯	◯	◯	◯	◯	◯	Strongly Agree

8. 7-I would imagine that most people would learn to use this system very quickly. *

Marcar apenas uma oval.

	1	2	3	4	5	6	7	
Stro	◯	◯	◯	◯	◯	◯	◯	Strongly Agree

9. 8-I found the system very cumbersome to use. *

Marcar apenas uma oval.

	1	2	3	4	5	6	7	
Stro	◯	◯	◯	◯	◯	◯	◯	Strongly Agree

10. 9-I felt very confident using the system. *

Marcar apenas uma oval.

	1	2	3	4	5	6	7	
Stro	◯	◯	◯	◯	◯	◯	◯	Strongly Agree

11. 10-I needed to learn a lot of things before I could get going with this system. *

Marcar apenas uma oval.

	1	2	3	4	5	6	7	
Stro	◯	◯	◯	◯	◯	◯	◯	Strongly Agree

8.8 Annex H

TQ-SOPI

Instructions:
We are interested in finding out what you feel about the experience you have just had in the 'DISPLAYED ENVIRONMENT'. We use the term 'displayed environment' here, and throughout this questionnaire, to refer to the film, video, computer game or virtual world that you have just encountered. Some of the questions refer to the 'CONTENT' of the displayed environment. By this we mean the story, scenes or events, or whatever you could see, hear, or sense happening within the displayed environment. The displayed environment and its content (including representations of people, animals, or cartoons, which we call 'CHARACTERS') are different from the 'REAL WORLD': the world you live in from day-to-day. Please refer back to this page if you are unsure about the meaning of any question.

There are two parts to this questionnaire, PART A and PART B. PART A asks about your thoughts and feelings once the displayed environment was over. PART B refers to your thoughts and feelings while you were experiencing the displayed environment. Please do not spend too much time on any one question. Your first response is usually the best. For each question, choose the answer CLOSEST to your own.

* Indica uma pergunta obrigatória

1. Please enter your number below: *

____

Group A

Please indicate **HOW MUCH YOU AGREE OR DISAGREE** with each of the following statements by circling just ONE of the numbers using the 5-point scale below.

1. Strongly disagree
2. Disagree
3. Neither agree nor disagree
4. Agree
5. Strongly agree

After my experience in the exhibited environment ...

2. 1-I felt disorientated... *

Marcar apenas uma oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly Agree

3. 2-I had a sense that I had returned from a journey... *

Marcar apenas uma oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly Agree

Group B

Please indicate **HOW MUCH YOU AGREE OR DISAGREE** with each of the following statements by circling just ONE of the numbers using the 5-point scale below.

1. Strongly disagree
2. Disagree
3. Neither agree nor disagree
4. Agree
5. Strongly agree

During my experience of the displayed enviroment...

4. 1-I felt myself being ‘drawn in’. *

Marcar apenas uma oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly agree

5. 2-I lost track of time. *

Marcar apenas uma oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly agree

6. 3-I felt I could interact with the displayed environment. *

Marcar apenas uma oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly agree

7. 4-The displayed environment seemed natural. *

Marcar apenas uma oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly agree

8. 5-It felt like the content was ‘live’. *

Marcar apenas uma oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly agree

9. 6-I felt that the characters and/or objects could almost touch me. *

Marcar apenas uma oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly agree

10. 7-I felt I was visiting the places in the displayed environment. *

Marcar apenas uma oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly agree

11. 8-I felt tired. *

Marcar apenas uma oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly agree

12. 9-The content seemed believable to me. *

Marcar apenas uma oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly agree

13. 10-I felt I wasn't just watching something. *

Marcar apenas uma oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly agree

14. 11-I had the sensation that I moved in response to parts of the displayed environment. *

Marcar apenas uma oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly agree

15. 12-I felt dizzy. *

Marcar apenas uma oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly agree

16. 13-I felt that the displayed environment was part of the real world. *

Marcar apenas uma oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly agree

17. 14-My experience was intense. *

Marcar apenas uma oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly agree

18. 15-I had a sense of being in the scenes displayed. *

Marcar apenas uma oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly agree

19. 16-I felt that I could move objects (in the displayed environment). *

Marcar apenas uma oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly agree

20. 17-The scenes depicted could really occur in the real world. *

Marcar apenas uma oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly agree

21. 18-I felt I had eyestrain. *

Marcar apenas uma oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly agree

22. 19-I could almost smell different features of the displayed environment. *

Marcar apenas uma oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly agree

23. 20-I had the sensation that the characters were aware of me. *

Marcar apenas uma oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly agree

24. 21-I felt surrounded by the displayed environment. *

Marcar apenas uma oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly agree

25. 22-I felt nauseous. *

Marcar apenas uma oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly agree

26. 23-I had a strong sense that the characters and objects were solid. *

Marcar apenas uma oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly agree

27. 24-I felt I could have reached out and touched things (in the displayed environment). *

Marcar apenas uma oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly agree

28. 25-I sensed that the temperature changed to match the scenes in the displayed environment. *

Marcar apenas uma oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly agree

29. 26-I responded emotionally *

Marcar apenas uma oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly agree

30. 27-I felt that all my senses were stimulated at the same time. *

Marcar apenas uma oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly agree

31. 28- I felt able to change the course of events in the displayed environment. *

Marcar apenas uma oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly agree

32. 29-I felt as though I was in the same space as the characters and/or objects. *

Marcar apenas uma oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly agree

33. 30-I had the sensation that parts of the displayed environment * (e.g. characters or objects)were responding to me.

Marcar apenas uma oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly agree

34. 31-It felt realistic to move things in the displayed environment. *

Marcar apenas uma oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly agree

35. 32-I felt I had a headache. *

Marcar apenas uma oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly agree

36. 33-I felt as though I was participating in the displayed environment. *

Marcar apenas uma oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly agree

37. If there is anything else you would like to add, please use the space below:

Este conteúdo não foi criado nem aprovado pela Google.

Google Formulários

8.9 Annex I

TQ-PQ

Characterize your experience in the environment, by marking an "X" in the appropriate box of the 7-point scale, in accordance with the question content and descriptive labels. Please consider the entire scale when making your responses, as the intermediate levels may apply. Answer the questions independently in the order that they appear. Do not skip questions or return to a previous question to change your answer.

* Indica uma pergunta obrigatória

1. Please enter your number below: *

WITH REGARD TO THE EXPERIENCED ENVIRONMENT

2. 1. How much were you able to control events? *

Marcar apenas uma oval.

	1	2	3	4	5	6	7	
NOT	◯	◯	◯	◯	◯	◯	◯	COMPLETELY

3. 2-How responsive was the environment to actions that you initiated (or performed)? *

Marcar apenas uma oval.

	1	2	3	4	5	6	7	
NOT	◯	◯	◯	◯	◯	◯	◯	COMPLETELY RESPONSIVE

4. 3. How natural did your interactions with the environment seem? *

Marcar apenas uma oval.

	1	2	3	4	5	6	7	
EXT	◯	◯	◯	◯	◯	◯	◯	COMPLETELY NATURAL

5. 4. How much did the visual aspects of the environment involve you? *

Marcar apenas uma oval.

	1	2	3	4	5	6	7	
NOT	◯	◯	◯	◯	◯	◯	◯	COMPLETELY

6. 5. How natural was the mechanism which controlled movement through the environment? *

Marcar apenas uma oval.

	1	2	3	4	5	6	7	
EXT	◯	◯	◯	◯	◯	◯	◯	COMPLETELY NATURAL

7. 6. How compelling was your sense of objects moving through space? *

Marcar apenas uma oval.

	1	2	3	4	5	6	7	
NOT	◯	◯	◯	◯	◯	◯	◯	VERY COMPELLING

8. 7. How compelling was your sense of objects moving through space? *

Marcar apenas uma oval.

	1	2	3	4	5	6	7	
NOT	◯	◯	◯	◯	◯	◯	◯	COMPLETELY NATURAL

9. 8. Were you able to anticipate what would happen next in response to the actions that
you performed? *

Marcar apenas uma oval.

	1	2	3	4	5	6	7	
NOT	◯	◯	◯	◯	◯	◯	◯	COMPLETELY

10. 9. How completely were you able to actively survey or search the environment using
vision? *

Marcar apenas uma oval.

	1	2	3	4	5	6	7	
NOT	◯	◯	◯	◯	◯	◯	◯	COMPLETELY

11. 10. How compelling was your sense of moving around inside the virtual environment? *

Marcar apenas uma oval.

	1	2	3	4	5	6	7	
NOT	◯	◯	◯	◯	◯	◯	◯	VERY COMPELLING

12. 11. How closely were you able to examine objects? *

Marcar apenas uma oval.

	1	2	3	4	5	6	7	
NOT	◯	◯	◯	◯	◯	◯	◯	VERY CLOSELY

13. 12. How well could you examine objects from multiple viewpoints? *

Marcar apenas uma oval.

	1	2	3	4	5	6	7	
NOT	◯	◯	◯	◯	◯	◯	◯	EXTENSIVELY

14. 13. How involved were you in the virtual environment experience? *

Marcar apenas uma oval.

	1	2	3	4	5	6	7	
NOT	◯	◯	◯	◯	◯	◯	◯	COMPLETELY ENGROSSED

15. 14. How much delay did you experience between your actions and expected outcomes? *

Marcar apenas uma oval.

	1	2	3	4	5	6	7	
NO I	◯	◯	◯	◯	◯	◯	◯	LONG DELAYS

16. 15. How quickly did you adjust to the virtual environment experience? *

Marcar apenas uma oval.

	1	2	3	4	5	6	7	
NOT	◯	◯	◯	◯	◯	◯	◯	LESS THAN

17. 16. How proficient in moving and interacting with the virtual environment did you feel at the end of the experience? *

Marcar apenas uma oval.

	1	2	3	4	5	6	7	
NOT	◯	◯	◯	◯	◯	◯	◯	VERY PROFICIENT

18. 17. How much did the visual display quality interfere or distract you from performing assigned tasks or required activities? *

Marcar apenas uma oval.

	1	2	3	4	5	6	7	
NOT	◯	◯	◯	◯	◯	◯	◯	PREVENT TASK PERFORMANCE

19. 18. How much did the control devices interfere with the performance of assigned tasks or with other activities? *

Marcar apenas uma oval.

	1	2	3	4	5	6	7	
NOT	◯	◯	◯	◯	◯	◯	◯	INTERFERED GREATLY

20. 19. How well could you concentrate on the assigned tasks or required activities rather than on the mechanisms used to perform those tasks or activities? *

Marcar apenas uma oval.

	1	2	3	4	5	6	7	
NOT	◯	◯	◯	◯	◯	◯	◯	COMPLETELY