UNIVERSIDAD COMPLUTENSE DE MADRID

FACULTAD DE INFORMÁTICA

TRABAJO DE FIN DE GRADO

Immersion and attention parametrization within 3D computer generated environments

Parametrización de inmersión y atención en entornos 3D generados por ordenador

---

Grado en Ingeniería Informática

Autores:

Fernando Muñoz Prado
Guillermo Monserrate Sánchez
Bruno Mayo Coronado

Directores:

Borja Manero
Alejandro Romero-Hernández

Curso académico 2021-2022

Special Thanks

We would like to take a moment to thank everyone that has supported us along the way, making possible the realization of an exciting investigation and development project that accentuates the most beautiful aspect of technology: how it can help people in their everyday lives.

Firstly, we would like to thank our tutors Borja Manero and Alejandro Romero, who have successfully guided us in every step there is to take in order to meet the standards in the field of project development. Their patience has proven to be as strong as their will to discover new horizons in human computer interactions, and for that, we can not be thankful enough. Along this year we have gathered so much more than mere technical experience, but the importance of investigation and groundwork in regards to human interactions, that leads to successful work, and settles the basis for any future work that may add exciting new aspects.

Of course, we can not forget all the support that was provided all the way from Barcelona's pedagogical team, the Didascalia team, more precisely, the notorious help of Ibis, Isabel, Miriela and Andy. Being a multidisciplinary project, it's essential to count with experienced professionals that will do anything in their power to transmit their knowledge to others. It is always an enriching experience to learn about topics outside of our usual professional scope (that is, psychology and pedagogy). We feel grateful to have shared this time working together on this project.

Lastly, we thank Universitàt Autónoma de Barcelona and Universidad Complutense de Madrid (specially the faculty of computer science) for making this project possible, and securing the future work that is to come. Also, we thank the support provided by Cognitive 3Ds developer team, who made possible the integration of 3D environment analysis tools within the Didascalia project.

Thank you all.

Abstract

Currently, despite the importance given to teacher training, there are not many tools or activities that help them practice in conflict situations that may arise in any classroom. Thanks to the Didascalias project, a virtual reality tool has been created so that teachers can practice these kind of cases by trying to resolve different situations. However, it has been proved that one of the most important problems generated by these situations is the general loss of attention on the part of the students when seeing an inexperienced teacher react to the conflicts.

Taking this into account, we have decided to take this virtual reality environment to a deeper level of analysis in this project, putting special emphasis on the participant's attention during the simulation in order to help them not to lose control on their students. The system captures the user's gaze at all times, showing all the points they have paid attention to, and for how long. This way, at the end of the execution, the user can be aware of where their attention was focused in each case.[1]

To test the effectiveness of the project, and the implementation of the new tool, we have performed a series of experiments followed by various improvements in each iteration. Thus, we have managed to collect data in the most efficient way possible until the last experiment, where we have concluded our final statements on the project.

The research allowed us to prove that most of the participants had not been aware of where their attention was during the test. In addition, thanks to the results of the professionals that had more experience, a better control of their attention was noticed, proving that the experience in this type of cases has helped them to acquire a better manage of the conflictive situation. Thanks to this study, and the opinion of the professionals that we consulted, we are positive that this tool can really help future teachers to improve their abilities in their educational formation.

Keywords: Education, gaze attention, non-verbal communication, virtual reality, classroom environment, teacher formation, serious games.

Contents

1 Introduction

1.1 Motivation

Studies have shown the lack of training for new teachers [2] may lead to students failing to learn the concepts imparted in the lessons and getting bad experiences which might end up making them want to drop out of school. By using a Virtual Reality serious game [3] and various software tools we will cover later, we wanted to help researchers and soon-to-be teachers understand some key concepts they might be doing wrong and correct it while giving them some experience to deal with difficult situations inside the classroom.

Motivated by the fact that this is a developing subject and the technology used to study it was very appealing to us, we decided to move on with it. Not only this but the opportunities that were presented to us, where we would be able to test the tool and the hardware with real life users and get to travel to Barcelona for the experiment and meet the research team there. All this together with the help of Alejandro and Borja kept us hooked to the project and motivated us to carry on with our work with them.

Perhaps one of the most beautiful aspects of technology, is its ability to improve people's quality of life on a regular basis, and on practically every field one can imagine. From the beginning of time, human progress has been severely determined by its technological advancements. It changes the way we think, we behave, how we interact with the world around us, and how we decide to shape our future.

This project emphasizes on the importance of high quality education, and the ways in which technology can provide an unparalleled environment for people to thrive and learn within conditions that would normally be quite hard to meet or even reproduce in real life. We desire to make a difference in regards to the future of education, and so, we present our humble contribution.

1.2 Didascalias Project

Aiming to help these new teachers, we developed some additions to an existing project, Didascalias. This Virtual Reality serious game was one of our main pillars to build this study and develop our ideas, by adding a new layer of complexity aimed at further analyzing the teacher's behaviour during stressful situations.

This project we joined in an advanced stage, allowed us to implement changes to their application and make it partly ours too. But most importantly, allowed us to experiment and make the best possible experience out of the project both in Madrid and Barcelona.

We also helped upgrading the already existing application, not only with the data collected for our work but also for enhancing the performance of the game, that had gone unnoticed for a great amount of time, and by taking into account the participant's feedback in order to provide an enjoyable and frustration free experience.

As a side effect, participants who found joy on the experience were also the ones that better performed when interpreting the roll of a "fake teacher" so to speak, inside the virtual environment the project provides, retroactively improving the quality of the experiments as well.

After all, its important not to forget the nature of the project, where people are at its center, and every effort is dedicated to capitalize on this aspect. It is to be considered a social work, putting people's quality of life as the very first priority, by setting up the basis for what the future of education could look like.

1.3 Objectives

As stated before, everything the project stands for (including our contributions), aims at improving the quality of learning experiences, always having people as the main character, so to speak. To achieve an extraordinary experience, there are several objectives that will set up the project for success.

Firstly, we made our main objective enabling the existing Didascalia project to able to track, study and analyze the attention patterns of its participants. We achieve this goal by closely examining one's non-verbal language when exposed to difficult situations that may arise from any educational environment (that is, frequent conflicts a teacher may encounter when lecturing).

Of course, such vast field can have many interpretations, so we had to focus on achievable goals for the scope of the development team. The parameter chosen as the most important for examining a participant's non-verbal language, is gaze. Where we look, for how long, and with what frequency we fixate on an object, are dead giveaways for information, including most importantly, data regarding our attention.

The secondary (but not less important) objective of our participation, is to highly improve the quality of the experience that teachers go through when participating on the experiments. Spending one's spare time freely taking part on a serious game that is laggy, unclear and frustrating, are not desirable conditions for anyone. By improving the player experience, we ensure that all participants have the best time possible when playing, and that they do not go as fast as possible to simply get over it. Enjoyment is part of a successful experiment.

This objective may derive on other, more specific sub-objectives, namely: Improve the game's performance to ensure a sufficient amount of fluency and as little loading times as possible, create an immersion for the player to forget momentarily about the real world and only focus on the virtual environment provided, and lastly, to guide the player in the most subtle and non-intrusive way possible (as to protect the immersion that we look towards to create and protect), providing only the strictly necessary information to successfully go through the game without interrupting the experiment to ask the organizers about some technical aspect of the game (for example, how to move, how to interact with the environment, when does a scenario end, etc).

1.4 Paper submission to the 30th International Conference on Computers in Education

We took all the aforementioned additions, and putted them to use with actual teachers. This comes in the form of presential experiments carried out in an effort to test how effective the system was on tracking, capturing and analyzing gaze data (along with player movement) and deriving from it how a participant's attention (or lack of it) can be perceived on the student's side, and how it can possibly affect in the class environment.

The full study regarding this aspect was submitted to the International Conference on Computers in Education (as a paper report format specified by the organizers), omitting the technical details involved on the implementation of software add-ons to make this parametrization possible, and focusing on the implications of non-verbal language and attention.

The paper is divided into an introductory section, where the motivations of the study are explained, followed with the previous project adaptation for the new 3D environmental tracking software tool, and it's evolution, following with details about the experimentation phase (including experimental design, participants, instruments, etc), and closing off with the results from the experiments and the conclusions derived from them.

As part of the conclusions, future work guidelines are specified to further emphasize on the level of parametrization that is theoretically possible to achieve. For example, it would be interesting to implement sensors for tracking vital constants, such as oxygen saturation and blood pressure, and other body language analysis aspects, more specifically facial expression analysis.

1.5 Structure

The structure of the document will be divided into nine sections. An introduction, followed by the state of the art, in which we describe the technologies involved in the making of the project, followed by an explanation regarding our work methodology. This includes the agile development used, the communications and coordination with Barcelona's team for pedagogical and psychological aspects to take into the account in the project development, and of course version control, as there are notable changes to the project's architecture.

After the working methodology, the evaluation phase will be explained, encapsulation the previous research done, the first experiments carried out, and the conclusions derived from them, as this will determine the work to be done on the next section: the application development phase. On this phase, the technical aspects regarding development are explained in detail. It encapsulates additions to the project, both on quality of experience and new functionalities, as well as a much needed performance upgrade.

After all these improvements take place, we describe the last experimentation phase, that in a similar manner to that of the first experiments, will define the future work to be done. The structure closes off with a section of final conclusions drawn across the year of experiences, the paper that was crafted to support the conclusions, and the bibliography and references.

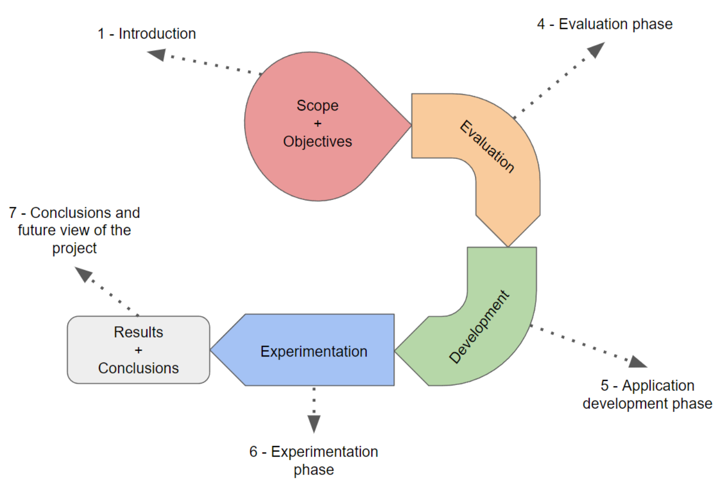

Figure 1: Lineout of the project phases and their corresponding sections

2 State of the art

2.1 Technologies and education

New technologies are developed everyday to help people in their job tasks, train to get some new ability or even help with health issues[4]. Education is no different from other areas and both teachers and students use some kind of technology daily in their routines[5]. Whether it is sending an e-mail, posting lessons in a virtual campus or teaching a lesson itself, technology takes a big role in education and it is increasing in importance as time goes on.

In addition the pandemic of Covid19 has affected billions of students who have had to change their lesson routines from attending to classes to fully remote or semi-remote schedules where they attend to video calls with teachers and have online lessons and meetings. Also the fact that this transition happened, or had to be performed so rapidly, made it so that the number of people having to use a computer, laptop or any kind of device to communicate with their teachers increased steeply during 2020 and 2021. This graphic shows the number of students who suffered this changes in their educational process.

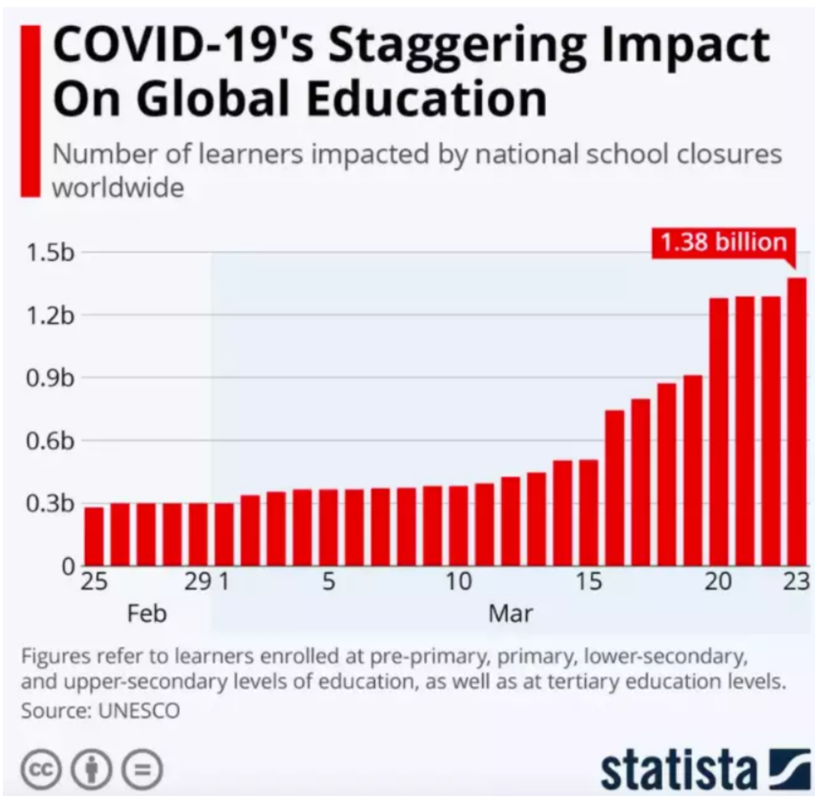

Figure 2: Students impacted by Covid19 [6]

As a result of these events, schools all over the world are investing more into technology to use it in the classrooms. This is also economically attractive to investors who are more prone to invest in companies dedicated to technology hardware and software for education [7].

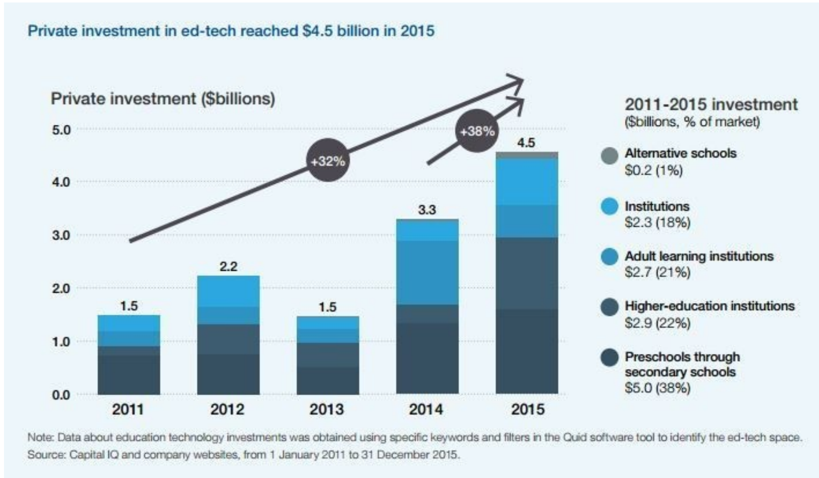

Figure 3: investment in education technology [8]

2.2 Video games applied to education

As one of these new technologies growing in the education area[9], video games are one of the main lines in these advances [10]. Because younger generations build up a lot of interest in video games during their free time, having them applied to education [11] makes it easier for them to learn new concepts while playing and makes the learning process lighter and funnier than traditional lessons. This is not only applied to younger generations as video games and simulators are used to train from airline and combat pilots to surgeons[12, 13].

In addition, the market for educational games has been growing since 2019 and it is forecasted to be more than double in 2030 according to [14]. This is a good sign as the growth in the market will surely produce more competition among companies who develop the games ending up in better and better products over the years and an even bigger growth in the sector. But most importantly, students will have better and easier access to these amazing tools for education.

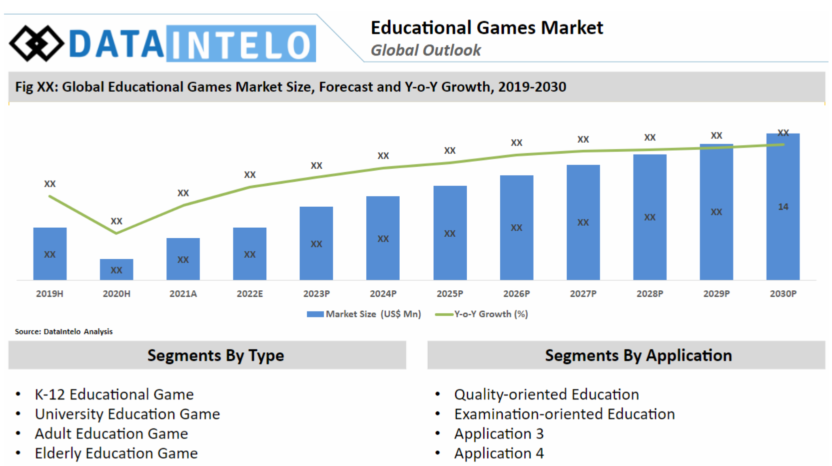

Figure 4: educational games market growth

2.3 Virtual reality applied to education

2.3.1 Importance of the immersive experience

When it comes to virtual reality, no matter what the purpose of the experience is, the immersion is a key factor to create an environment that is as close to real life as possible [15]. Using a Head Mounted Display is not something we do everyday and a poor developed experience may result in false data (in our experiments for example) or people not having fun or even having dizziness issues while using the VR goggles.[16] This level of immersion is accomplished by paying attention to the smaller details, which combined are capable of delivering a realistic experience. These include sound, visual details such as furniture and realistic lights in interior models and even birds flying in the distance in a nature-like environment. In our case the classroom is laid out as it is in different high schools across Spain.

2.3.2 3D environment generation

The generation of these 3D environments is performed using different 3D modeling tools such as Blender, Autodesk Maya or Cinema 4D among many others. Unity's development platform allows the designers and programmers to import these environments to a project and use them as the scenery for their video games whether they use virtual reality or not. We used these tools and imported the resulting unity project into our data collection tools, which we will explain further in this document.

2.3.3 Environment simulation

For us to be able to translate what happens during the experience to some data that we can process afterwards, we need some tools to replicate the 3D environment and track what is happening in it. In order to accomplish this, we made use of Fove SDK even though this is not the only one in the market, it is the one we used because it is compatible with older VR headset models and with the Unity version the project is developed in.

2.3.4 User's vision tracking (gaze)

In order to improve the functionality of VR headsets and get additional data on how the user interacts with the environment, gaze tracking tools are being developed among other data collection software. These features are key in order to properly track what the user does since usual eye tracking tools (IR trackers for example) are difficult or impossible as they require extra hardware that is incompatible with the head mounted display [17]. Software solutions are a better option in this matter and a better fit for our project since the cost of eye tracking hardware is very high since the devices have to be specifically made to be used with your goggles of choice and can cause motion sickness and headaches to the user in some cases. In the other hand, software such as TobiiAR, Cognitive3D and some others offer not only the gaze tracking functionality but also an interface or dashboard where the collected information is clearly displayed. The latter was the best choice for our project[18].

2.3.5 User's movement tracking

In addition to tracking the gaze and the focus of the user, it is also relevant to track the movement inside the 3D environment as well as the body language of the subject [19] while taking part in the experience. for tracking the body movement, the most effective and common tool is a motion capture suit (Figure 1) which tracks spots of the body with sensors and records and recreates the movement [20]. This suit together with software such as Emopose, allows us to get data about the feelings of the person being recorded. Emopose is able to detect different sentiments depending on the movement it captures. In the other hand, data collection tools such as Cognitive3D allow us to get a **heat map** of the areas more visited by the user and thus track their position in the 3D environment during the experience.[21]

Figure 5: Motion capture suit

2.3.6 Comparison among users

In addition to all the points mentioned before, we shall remark the fact that different users have different levels of expertise in the use of virtual reality headsets and different levels of experience in the matter of our study and this must be taken into account when extracting the results from the experiments. Some inexperienced users may feel dizzy at first or may not know how to move around the environment. As this could lead to wrong data, the creation of a guide within the application is really useful to get rid of all possible misunderstandings and having someone supervising the experience is also useful in this regard.

2.4 Video game data collecting tools

In the serious games field, giving technical data and analysis is crucial for the success of the experiments and the usefulness of the studies developed [22]. Not only in this field but also for upgrading the performance of the applications and providing a better user experience, it is very important to go through an investigation phase where system data while execution of the application. This will give the software designers and engineers a good starting point to make better updates to an existing software [23]. Data collecting tools are key in these processes and usually come built-in with other applications.

2.5 Video game development tools

There are numerous development environments and tools to create not only video games and program in the different programming languages. The choice you make when it comes to the environment you use comes down to personal preference and the language and APIs depend on what kind of game or application you are developing. In our case, we developed our project using C# and Unity's environment as well as Unity engine.

2.5.1 Game engines

Game engines consist on a series of tools gathered in an environment that is designed to create video games. They usually include an editor where the developer can add and remove objects in 3D and 2D scenes, apply all kinds of parameters and settings, creating menus, managing audio and building the different blocks of the game. They also have a physics engine that consist of a software that provides a base to give velocity, mass and all the different physical properties to the objects of the game and tune them to your preferences to be able to create anything from scratch. For example you may want to create space like feeling of zero gravity or you might need your character to have super speed or be super strong. This is achieved tweaking the physics applied to the character and the objects of the world we are creating.

- **Unity**

  Unity is one of the biggest game engines in the industry, not only because it has been developed and upgraded over a long time but also because it is virtually free for a non commercial use. This has allowed a lot of investigation projects to be hosted inside unity's environment [24]. Unity provides a working space where everything is laid out for the developers to have an easy work-flow and ease to add new objects, scenes and scripts to the game. It also has a big asset store where artists post their creations and sell them or give them out for free to the programmers and designers. Scripts for Unity projects are written in C# programming language although other scripts in different languages can be launched for tasks simultaneous to the game execution such as data collection for further analysis.

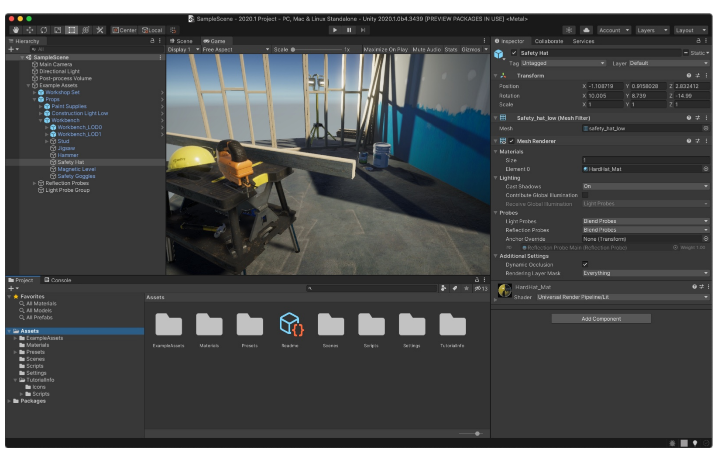

Figure 6: Unity developing environment

- **C**

  C# is a programming language developed and standarized by Microsoft as part of .NET platform. It's syntax is derived from C/C++ syntax. It was invented to program applications in .NET but it is an independent programming language and it can be compiled for different operating systems. This makes it a suitable option for multi-platform applications and a good option for video games because developers only need to program one single game and put it out in different platforms as oppose to building one version for every different operating system, game console etc [25].

2.5.2 Human language analysis

- **Vokaturi**

  Vokaturi is a speech recognition tool that analyzes the speech in real time and in post-execution by inputting the recorded speech into deep neural networks and gives a score for every detected emotion expressed during the recording time. It can also be trained to measure some specific emotions or specific patterns.

- **VADER Sentiment**

  VADER sentiment analysis tool (Valence Aware Dictionary and Sentiment Reasoner) is an open source software designed to detect the sentiments expressed in social media texts. It also works with any other kind of text. It is based on different rules and the lexicon used in the sentences. Taking words individually and asigning how "good" or "bad" they are, the tool is able to classify the degree of different sentiments expressed in the text.

- **SpeechRecognition**

  SpeechRecognition is a python library that allows to listen to audio through the user's microphone and turn it into plain text. This is useful combined with VADER sentiment analysis tool to provide voice analysis of an audio fragment that can be captured while the execution of another application (in this case a video game).

2.5.3 Movement and eye tracking tools

For our project, it was highly important to have good vision and motion capture technology [26]. After investigating several alternatives, we decided to opt for Cognitive3D, due to its high capacity when it comes to representing areas affected by vision, as well as all the tools it has, such as monitoring dynamic objects.

- **Fove integration**

  One of the best implications of Cognitive was the ability to use FOVE, the first Virtual Reality Head-mounted Display that uses eye tracking. Thanks to this tool, we were able to carry out the core of our study, monitoring the user's gaze at all times.

- **Dashboard**

  The dashboard allows you to have control of all the data that can be extracted from a session. Once the scenes are uploaded, it allows you to enter each of the possible configurations. Among

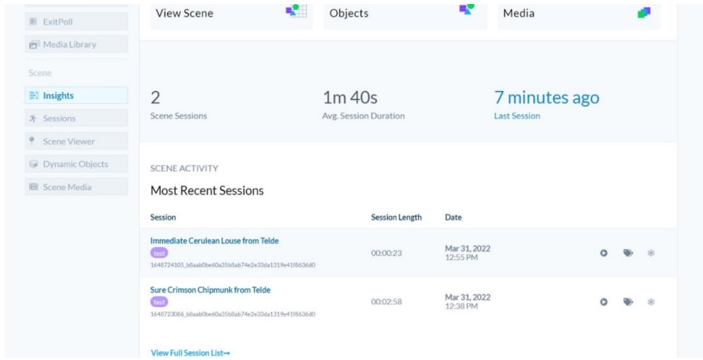

Figure 7: Main visualization of the dashboard from Cognitive3D.

the tools to take into account, there are those that we are going to give more importance to in the project.

Among them we find the scenes, the main responsible for everything being in order to collect the data in the experiment. These scenes work like a Unity scene, uploading the geometry of a project's scene to the dashboard. Once the experiment starts, the session starts collecting all the data using that scene as the main scenario.

It is also important to talk about dynamic objects, in this case the students. Their existence will allow Cognitive3D to interact with them when extracting data, especially taking into account their position. For example, if a student moves in one of the scenarios, the session recognizes that he is moving and records his entire path. In this way, when the user follows the student with his eyes, the space that he has occupied will be reflected at all times.

Within these sessions we have different ways of visualizing the data. In the active session view, we can see the entire execution in real time, showing the user's gaze, and reflecting it in a heatmap that shows the parts to which he has paid the most attention. There are also other forms of visualization, such as comparisons between several sessions, or a map where the gaze is reflected through cubes. We will talk about these tools later, when we present the data collected in the experiments.

In addition, the dashboard allows us to give each session different labels. In this way we can keep each of the sessions well differentiated, referring to each scenario and participant. This functionality will be key when comparing data between participants.

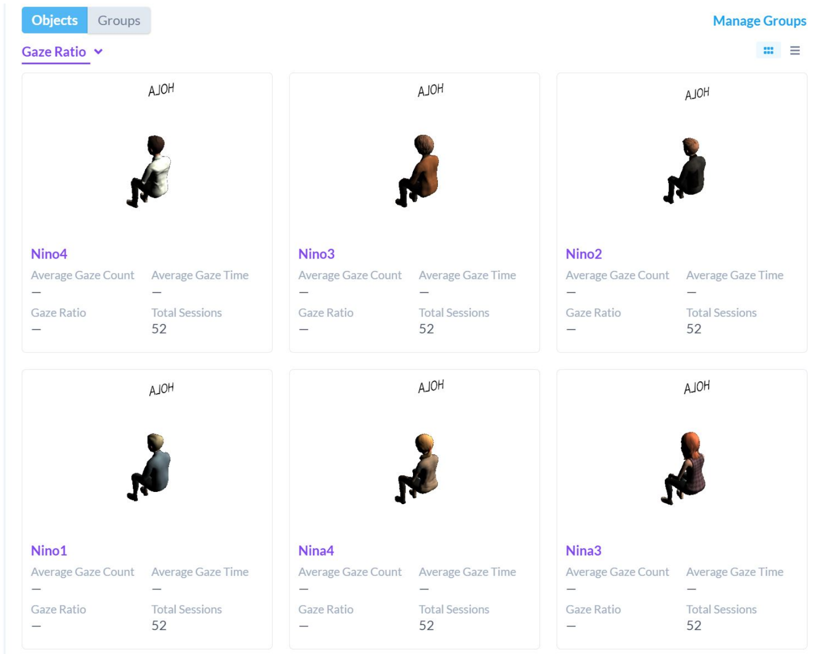

Figure 8: Display of the dynamic objects in the scene

Figure 9: A session with different tags to identify it.

2.5.4 Virtual Reality Hardware

For this project, we used the Oculus Rift S virtual reality glasses, which were given to us by the Computer Science Faculty of the UCM. These VR glasses allowed us both to test the development part of the tool, as well as to allow the participants of the experiments to have the necessary material to carry out the test.

For the experiments in Barcelona, we also had at our disposal another VR glasses, the Oculus Quest 2, provided by the UAM team. This advantage allowed us to carry out the experiments simultaneously, therefore we were able to do many more in less time, collecting even more data.

Although the Oculus Rift S is not the newest model, it has a very well built base and therefore a very low level of failure. This makes them one of the most stable VR tools for experimenting. Among its features, the ones that are most important for our project are the following:

- **Sound**: the device has speakers built into the glasses themselves. In addition, it also has a microphone that is part of the device. These two characteristics are very useful to us when carrying out the experiments, since we can hear what the participant is hearing, and see if there is any error. In addition, the integrated microphone means that an external microphone does not have to be placed, making the experience much more comfortable.

- **Display**: the visuals are better than its predecessor, with a 1280x1440 per eye. It is also noticeable that it has an LCD display. This way, the perceived screen door effect is very stable, which is good for our project.
- **Controllers**: the controllers of the Oculus Rift S are very ergonomic. They consist of two buttons, two triggers, a thumbstick joystick and a system menu button.
- **Oculus Insight**: this feature that allows movements to be translated in virtual reality regardless of the user's position. It offers room-wide tracking without the need for external sensors to pick up user motion.

Figure 10: Components of the Oculus Rift S virtual reality glasses

2.6 Didascalias

2.6.1 What is the project about?

Didascalias consists of a research, development and innovation project, based on educational propouses. The project is made up of a team from different branches, such as education, pedagogy, psychology and engineering. In our case, we have worked as the technical part of engineering (from the Universidad Complutense de Madrid), working hand in hand with the sociological and pedagogical part of the field of education (from the Universidad Autónoma de Barcelona).[**?**]

Its main objective is to facilitate the training of future teachers, creating a virtual reality environment through which both their verbal actions and their non-verbal language can be analyzed [27], in order to improve their communication skills for the class and their students. More specifically, it puts teachers in conflictive situations to see how they resolve them, since these types of situations are the most enlightening when it comes to analyzing them.

We thought this would be a very good starting point to develop our study since body language takes a big part in these situations where the teacher needs to keep his/her authority for the lesson to run smoothly [28] [**?**]. Starting from the Classroom VR project belonging to Didascalia, we have decided to take it further by analyzing non-verbal language more specifically. Therefore, we began to analyze the teacher's attention in the classroom, monitoring his gaze at all times. In this way, we seek to achieve a larger scope of the attention paid by the user throughout the execution, emphasizing the conflictive situation.

Figure 11: Execution of the third scenario.

2.6.2 Importance of facing conflicts in the classroom as a teacher

The good skills of teachers are essential to carry out a good educational system. A teacher served by good educational training is essential when transmitting information in the classroom to the students [29]. However, it is shown that the vast majority of teachers do not receive any training in order to assimilate how to manage a classroom of students [30], and it is mostly because it is declared that they do not have access to this type of training.

There are already many studies and projects that speak of the importance of good teacher training using virtual reality environments to achieve this. In 2017 a study was carried out at the Ktisis Cyprus University of Technology [31] to show the potential of virtual reality to be used as a learning tool for teachers. Two experiments were carried out, the first consisting of teachers trying to detect and understand possible disorders of the students in the classroom, such as vision problems. The second focused on the ability of teachers to detect cases of bullying within the classroom environment. The final results showed that the VR experience helped teachers better understand the students by putting themselves more easily in their position. Based on this study, new iterations of the application have been carried out with the aim of increasingly improving its effectiveness.

Another project worth mentioning is that of VR Teacher [32], a Virtual Reality based training to improve digital competences of teachers in crisis situations. This project presents a virtual classroom where the teacher can acquire skills and experience through scenarios that are given during a crisis, such as a pandemic.

Different approaches have also been made to the study of teachers' attention and gaze, such as [**?**], in which they used an eye tracker, not as an app but as a video. In this case, they brought together both student teachers and teacher educators, to whom they showed a video of a real incident generated in a class, with the intention of asking them later what had happened. The data showed that experienced teachers were able to better understand the problem, and found that eye-tracking data can help identify a professional view within the incident.

Therefore, the importance of improving the skills of teachers in the classroom is high, and it is necessary to find an option that makes up for the limitations that are found in education at this time.

2.6.3 Combined work with UAB

As previously introduced, the Didascalia project is made up of a team from various sectors. During our work, we have cooperated hand in hand with the team of the Universidad Autónoma de Barcelona to be able to work in parallel and as efficiently as possible.

First we started working remotely until we became devirtualized in the first experiment, carried out in the I.E.S Carmen Conde. During this experiment we had the opportunity to get to know them better and soak up their knowledge and skills. Later, we had the opportunity to work with them again in the experiment carried out at the Autonomous University of Barcelona, where we were able to meet more of the team and carry out an even larger two-day experiment.

After these two experiments, we have kept in touch with the team as we developed our most important part of the project. Thanks to everything we learned in the previous meetings with the team, we were able to carry out a third and final experiment at the Colegio Ramón y Cajal completely managed by us, putting into practice everything learned during the previous meetings, but always taking into account the help of the entire team.

3 Working methodology

3.1 Agile development

Throughout the project we have met constantly and periodically, adopting an agile methodology as the main structure, due to its speed, dynamism and efficiency.

One of the main reasons for choosing an agile methodology, apart from the fact that it is fast and dynamic, was that the client is considered as part of the production team. In our case, the client could be considered the Didascalia team, since our job is to improve this tool. Therefore, the team was always aware of all the modifications and ideas that were coming up.

Within the agile methodologies, we chose to start with DevOps, a methodology that is focused on software development and ensures that the launch of ideas is constant. Choosing this method at the beginning of the project was very successful, since in this first part we were looking to get to know the application that already existed, as well as start launching ideas to find one on which to focus our project.

Once we decided on the main idea of our project, we moved on to the Scrum method, leaving aside the constant launch of new ideas and focusing on the one we had decided on. In this way, we were able to maintain an agile methodology, but more based on the objectives we had in mind for the end of the project.

Finally, when reviewing the project we have been able to see several advantages of having made these decisions reflected in the final results. Among them we highlight:

3.2 Communication with the external team in Barcelona

- **Once the main priorities are decided, resources and results are optimized.** This point is very much reflected on our work once we decided to focus on the study of attention in the gaze, since we had a list of tasks that we were touching on, but we immediately unified as a team for the immediate priority: to implement Cognitive3D in the project.
- **The client (Didascalia) can contribute their opinions so that the project constantly improves.** In this case, we always had the constant help of the Barcelona team, which meant constant verification of our work. This also prevented the work we had already done from having to be redone, since we made sure we were always on the right track.
- **The ability to use solutions during the work process.** This point was key, since we did not have to wait long periods of time to be supervised. We were able to employ solutions without fear, as we knew that it would be verified soon and could be rectified if necessary.

3.3 Version Control

For version control we use our own repositories, in order to keep track of all changes, Github and Google Drive. Although at the beginning we made all the changes together, later on we had to branch out into various tasks. Therefore, maintaining version control was essential. First we had a single branch of the project. However, during the process of introducing Cognitive3D we created a new branch, since we did not know if its incorporation was going to be definitive.

Also, Google Drive was very helpful in planning many features of the project, not just the Unity code and files issue. In this unit we made a division by folders that allowed us to constantly monitor all the documentation as well as planning files, research and other documents of interest. We made the following division:

- **Memory:** all documents related to the final report, such as the structure of the index, reference links, articles of interest or images and tables to be used
- **Project Version:** to bring the best track of each version, and have them on hand not only for us, but also in case the Barcelona team needed it, since it is more accessible to them than Github.
- **Research:** all the documents related to the research, both the first ones when using the tool, and the individual ones when we start looking for alternatives and ways to improve the project
- **Captures:** this folder refers to the screen recordings that we were making in all the experiments, at the request of the Barcelona team, which were very useful to compare the executions before implementing Cognitive3D.
- **Own code:** in this folder we save the external code contributions that were made outside of Unity
- **Paper:** all the documents related to the paper that we carry out on the research.
- **Performance:** this folder is dedicated to one of our goals in the project, the complete optimization of it to give a better user experience.

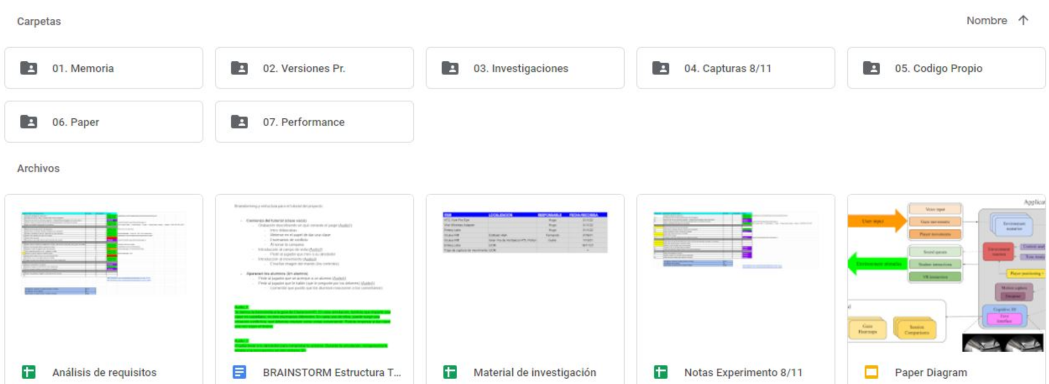

Figure 12: Organization of the Google Drive Unit of the project.

4 Evaluation phase

The first step on the journey for every development process, begins with an exhaustive research around the existing product, and, given the nature of investigation projects, there needs to be a well defined plan of action. We may consider splitting the evaluation phase into the Investigation, and Experimentation phase.

In the investigation phase, we define the information gathered from documentation, meetings, and code analysis. In the experimentation phase, we describe the first approach taken towards practical experimentation (that is, the first of several rounds of experiments, where no remarkable changes have been applied yet).

4.1 Research

This phase consists primarily of getting familiar with both the code and the structure of the project. To achieve this goal we followed a top down approach throughout the project's directories, noting down which might be the most important files for future modifications (this discards media resources, Unity asset files, third party software add-ons that are not meant to be modified, and various other components, like modules for hardware interactions).

Investigation begins with reverse engineering from the pillars of the program logic, logging every detail that may have relevance when trying to implement new features, modify the flow or correct known errors. After a few weeks, we had on our hands a list of questions to be touched down on at the meeting with one of the developers of the previous project.

4.1.1 Training with previous developers

Training with the former developer was arranged in an effort to further deepen the understanding of the intricacies regarding the project, starting with the aforementioned list.

Some of the key aspects discussed at the meeting were: general organization and structure of the project directories and the reasoning behind it, potential tweaks to improve the quality of the experience, awareness on dangerous areas to modify and potential scope of development.

This last theme includes the possible contributions to be made (aside from the quality of product changes mentioned before, which mostly constitute low magnitude changes), taking into consideration the previous efforts made during the various iterations of the project.

4.1.2 Study of the flow of execution

Understanding the flow of the program is essential when trying to implement future changes. As will be described later, the flow will experience some minor modifications in order to accommodate a new relevant usability approach. The existing application flow describes as follows:

Firstly, the player is greeted with a main menu, providing access to either the program termination, or the secondary menu listing the possible scenarios to play. Inside each scenario, after a certain period of time, the key action will be triggered, followed by an interrupt meant to give the player time to think through his/her possible course of action to solve the problematic situation.

To reactivate the flow, there needs to be a key press by someone other than the player (that will have the virtual reality headset on) and after the player has made its move (again, after a certain time frame), he/she will be given auditory feedback explaining the correctness of the action taken. After the termination of each scenario, the execution stops (meaning that looping through all three scenarios requires three separate program executions).

4.1.3 Dependencies

On this section, the most relevant C# files will be briefly described. These files represent the core functionality of the project, managing most fundamental logic of the code, and therefore almost every other script has some degree of dependency over them.

GameManager is responsible for starting the game as a whole. Contains parameters that control the display type of the execution (either standard of VR), hosts static attributes containing object instances that will be needed during the execution of the scenes (most notably MySceneManager), and has controllers moving across the game scenes, whether it is the main menu, secondary menu or through the different scenes.

MySceneManager is responsible for the common logic shared between the game scenes. This includes, among others, controlling the triggers to fire actions inside the scene (including when to start and finish scene execution), handling the user input (both direct like for example, voice input, and indirect, like actions carried out according to the player's movement around the class), initializing and controlling the actions of the student objects (positioning, problematics, animations, etc), sound queues and actions referring to the different path variations according to the user's choices during each scene.

The unique characteristics of each scenario are defined in the directory SceneScriptables, found in Assets/Scripts. Scene 1 and 2 share a scene scriptable file (as the set up needed is similar) while scene 3 has its own. There are also scriptable object scripts, that define properties of the game objects (such as the class morphology information, path information, etc) and other more specific managers, like MotionCaptureManager (with the purpose of acting like an interface between the program and the motion capture suit) and SoundManager, that controls audio input and output (the audio input is used as a controller for some actions during the scenes).

4.1.4 Initial requirement analysis

Closing off the investigation phase, we define an incremental list of requirements that need to be addressed in the early stages of development. These requirements mainly focus on tweaking details about the player experience and carrying out the necessary preparations for future stages of the development.

Tasks are divided into different priority groups that will determine when the task is to be completed, and the estimated complexity that a task may possess (being a proportional relation with the required time to achieve them). Note that, for each task, there is a status cell, indicating the current situation of the task. As time progresses this state will change according to their completion.

Also, some tasks may be stalled indefinitely if a higher priority objective arises outside the initial requirement consideration (that is, for example, greater scale additions, further discussed in detail on the "Application development phase").

#	Description of the task	Priority	Complexity	State
1	Extend scenario 1 time to trigger criticallity	1	4	COMPLETED
2	Guide for making the program work + trouble shooting + VR setup	1	3	COMPLETED
3	Continuos execution of the program + player independance	1	1	STALLED
4	Eliminate the options given before each scenario to avoid biasing the participants	1	2	COMPLETED
5	Adjust player movement sensitivity	1	3	COMPLETED
7	Add a hover effect to better select each menu option	2		COMPLETED
8	Take out the step on the floor because it limits player movement	2	4	COMPLETED
9	Eliminate the feedback given after each scenario to avoid biasing the participants	2	2	COMPLETED
10	Add extra names to avoid repetition	2	5	COMPLETED
11	Add pause button with VR controller	2	3	
12	Set problematic student with normal text color	2	3	STALLED
14	Add soundqueues for the end and start of each scenario	1	3	COMPLETED
15	Add chatter to the start of the class scenarios	3	3	COMPLETED
16	Register player path	2	2	WIP
17	Create emoption analysis script and graph output			COMPLETED
	Take out the pause after criticallity			COMPLETED
18	Implement tone script on final project solution			
19	Ingame guide to introduce the players to the experience	3	2	COMPLETED
20	Get itial base tone for better voice analysis	3	?	COMPLETED
22	Add blackboard functionallity	4	?	STALLED
23	Extend the movement space of the class through chairs and tables	4	3	COMPLETED
24	Add class decoration to better represent a real class environment	5	5	STALLED

Figure 13: List of requirements analysis for tasks and statuses

4.2 Experimentation

To finish off the evaluation phase, there needed to be an empirical part to confirm our prior assumptions. In this section we will describe the process that follows from extensive code analysis. The main focus being, to test on a real life setting the project's capabilities. As expected, fieldwork shed light on non obvious problems that the program sustained, as well as suggesting different areas that could benefit from polishing, or entirely receive a different approach to accomodate a more ambitious take on parametrization of human-computer interaction, specially attention studies through player gaze focus.

4.2.1 Coordination with the team in Barcelona

Of course, being a project built around the learning experience of novice teachers, there needs to be strong guidance by the hand of professionals of the field, as was made in the previous years of development. For this reason, periodical meetings with the pedagogical team of experts, over at the Universidad Autónoma de Barcelona were arranged, in order to plan out and materialize the future development of the project. The meetings were fundamentally focused on discussing the new tasks to be added to the aforementioned requirement list. Tasks were carefully picked, always having in mind: to not influence the user when making decisions during the game, to provide exactly the necessary information to complete the experience (no more, and no less), and to make the overall experience as enjoyable and discomfort free as possible.

4.2.2 User's manual for UAB team

Using a software product while it is still in production can prove to be quite troublesome, especially for people without the technical background a project developer would have. In addition, the pedagogical team over at Barcelona would most likely need to conduct several experiments without counting with the support of the development team.

For these reasons we made a guide that condensed what we thought to be the most important aspects needed to acquire a certain degree of independence when running project experiments. The guide included an introductory section stating the technologies needed, a run down on how to execute the Unity project from within it's editor, the virtual reality headset configuration guide, and general scenarios for troubleshooting, among which we can point out: headset connectivity issues, controller issues, computer thermal throttling, Unity crashes, version control and memory corruption, to name a few.

4.2.3 Experimental design

Though with variant arrays of objectives, the experimental design across the different empirical experiences remained fairly constant. The experiments would always take place in educational institutions that kindly decided to collaborate in the project. There, any teacher (regardless of the background or area of expertise) could anonymously take part in the experience, which would consist of going through three conflict scenarios that may arise in a standard classroom environment, always reacting as they deemed necessary, and completing a short questionnaire and interview afterwards. Both the game experience and the interview were recorded as testimonial back up for possible future reference.

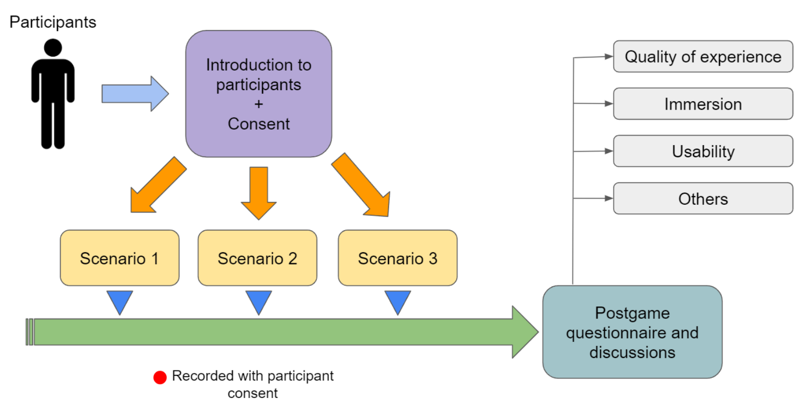

Figure 14: Visual description of inital experimental design

Note that, in latter stages of experimentation, the contents of the questionnaire shift towards aspects regarding software product quality, user experience, future adaptability and attention study, rather than the pedagogical focus these first experiments are bound to have. This is because as the project advances, some other areas of development require fieldwork testing to prove their effectiveness and further define future goals.

4.2.4 Experiments of the First phase

Previous to the application development phase, there were two instances of presential experiments conducted on the newest version of the game at that time, and counting with the help of Barcelona's pedagogical team, as was made on the previous years of the project's development

- Experiment in Madrid (I)

  This was the very first experiment conducted by the current development team. It took place at the institute I.E.S Carmen Conde on the eighth of November 2021. Process was executed as stated before, having 14 participants voluntarily take part in the experiment, all with considerable years of experience in the teaching field (although in different areas of secondary school subjects). For each scenario, the game classifies the action taken as appropriate or non appropriate, and provides auditory feedback accordingly. Having this in mind, the results read as follows (the implications regarding these results will be discussed in conclusions following the experimental phase):

Figure 15: Participant on Madrid's first experiment.

- Experiment in Barcelona

  The experiment on Barcelona's Universidad Autonoma was carried out in a similar manner to that of Madrid's, with a few differences:

  The experiment execution was splitted between two days, as the sample space was much greater, counting with several entire class groups of participants coming from the master of pedagogy. Being this the case, it is obvious that in this experiment, participants did not count with experience lecturing in the classroom, which is equally interesting when considering different study groups may provide more diverse insights on the project.

  Given the high volume of participants, they were told to form groups, of which one student would go through all three of the virtual reality scenarios, undergo the interview, and discuss with teammates the experience, forming what is commonly known as a focus group.

  Once the conclusions of each focus group were defined, these were sent to the heads of the pedagogical experts team, who then provided a curated file containing all relevant information towards the project development.

Scenario	Times appropriate ending was achieved	Times non appropriate ending was achieved	Times the scenario timed out with no result
#1 - Inappropriate comment	1	12	1
#2 - Problematic union	1	13	0
#3 - Group dynamic problem	4	5	5

Table 1: Experiment results

Figure 16: Participant on Barcelona's experiment.

4.3 Conclusions following the first experimental phase

Following the data obtained by means of participant interviews and informal discussions, we arrive at important conclusions that will define the development phase's priorities and objectives to break through.

Without a doubt the most important conclusion obtained from the experimentation phase was the significance of immersion for virtual reality experiences [33]. Numerous participants stated that while being mostly acceptable, the environment did not quite feel like a real classroom, mostly due to ambiance lighting, lack of common noises that one would find on such environments (like for example, light chatter, beginning and ending ring bells, etc) and low motivation to achieve the desired level of participation by the player within the environment. This combined with the fact that many participants did not know how to react, when to do so, or to what extent they could interact with the emulated students, manifests the necessity to create a new scenario previous to the actual experience. That is to say, an interactive guide that explains the key aspects of the game in the most non intrusive way possible, while motivating the player to actively participate through the experience, improving it and exploiting the capabilities of the game.

Also, when the 3D analysis software is implemented for studying the participant's attention, the guide will cleanse almost all possible biased behavior that results from the inexperience of the common user getting introduce for the first time to virtual reality, so that the data collected regarding attention based on gaze [34], is valuable because it closely represents the behavior that the participant would have on a real life scenario. Immersion is also extremely relevant in this aspect because there exists a direct correlation between attention and immersion: the more immersed a player is on the game setting, the more likely she/he is to focus solely on the virtual environment and momentarily forget about the real world environment, providing excellent quality data that closely resemblances the one to be obtained in real life.

Performance is perhaps the great forgotten when developing software projects. Some of the participants stated experiencing some kind of discomfort that can be traced back to performance issues. Details mentioned by participants include low framerates, various instances of screen tear and input lag, and heavy loading times between scenarios. Also, it should be noted that as the experiment execution went on, there was considerable thermal throttling (that is, when computer components overheat so much that they have to actively lower their performance to protect themselves) resulting in even lower performance over repeated executions of the game, making this, an important matter to solve on the development phase.

Finally, as in any software project, multiple bugs emerged. Error correction will be handled inside the requirement list mentioned in the "Initial requirement analysis" section, as a part of general quality of life changes, aimed at improving the user's quality of experience.

5 Application development phase

Once the requirements and goals to be met are well defined and studied, we proceed to the development phase. Here, the exact technical procedures will be described and documented in detail. The section will be splitted into contributions, that groups up new features implemented into the project, and performance upgrade, that groups all the tasks aimed at improving the program's hardware requirements and polishing the user experience by eliminating the mentioned discomforts traced back to performance issues.

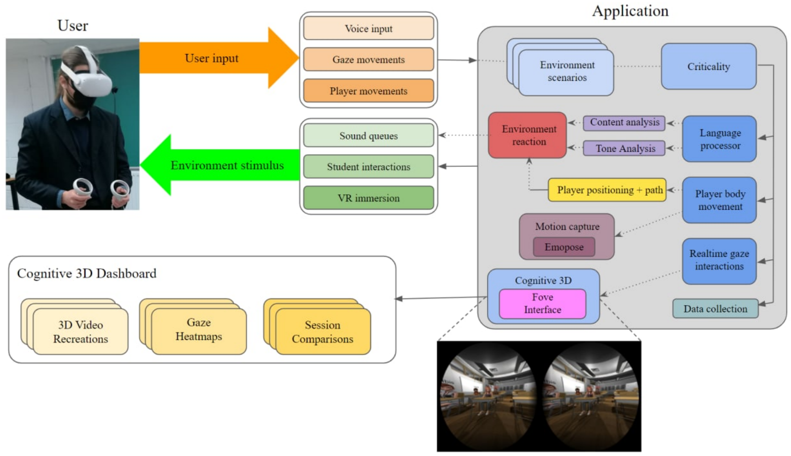

Figure 17: Program's information flow graphic description.

5.1 Additions to the project

These are the relevant additions and improvements carried out both as an effort to improve the overall quality of the project, and to provide new functionalities focused around immersion and parametrization of the player's body language (especially gaze), as well as setting an adequate user experience to further capitalize on player behavior parameterization in the future.

5.1.1 New functionalities and general modifications

Following the participant's advice, the first modification made was the adjustment of timers for the critical events to take place on all scenarios, as being too short resulted in the participants not being able to fully form an artificial lecture to get going on the simulation. The list of the students' names was also increased in order to create the impression of a wider student group (before, many names would constantly repeat making it difficult for the users to refer to a certain individual in the classroom). Also, we added a simple hover effect on the menus so the item selection would be much more readable, as selecting an item within a menu without noticing any change is quite confusing.

There were several other changes to be made for achieving better user experience. The first modification to be made on this matter, was the removal of the class' pallet (that is to say, the extra step found on the teachers' side of the class), as many participants stated feeling awkward when taking a ghost step that really didn't change the player's vision hight, resulting in a dissonance between the expected height drop and the actual behavior of the program's physics.

Similarly, the teacher's desk was also removed, following the advice from Barcelona's team, as they suggested that having a solid object in front of the participants would unconsciously make them more susceptible to staying idle instead of moving around the class while lecturing, as was intended in the first place. This change, combined with the removal of the pallet, resulted in more instances of natural player movement around the environment.

The last modifications on this section were aimed at capitalizing on the previous improvement on player movement. As the majority of the participants found trouble controlling the movement of their character, we decided to smoothen the movement transferred by the joystick; this was achieved by removing the speed acceleration and reducing the sensibility on the headset controller. Also, as the player's collider was interfering with that of the classroom's table mesh collider, we had to both adjust the model collider width and spatial positioning.

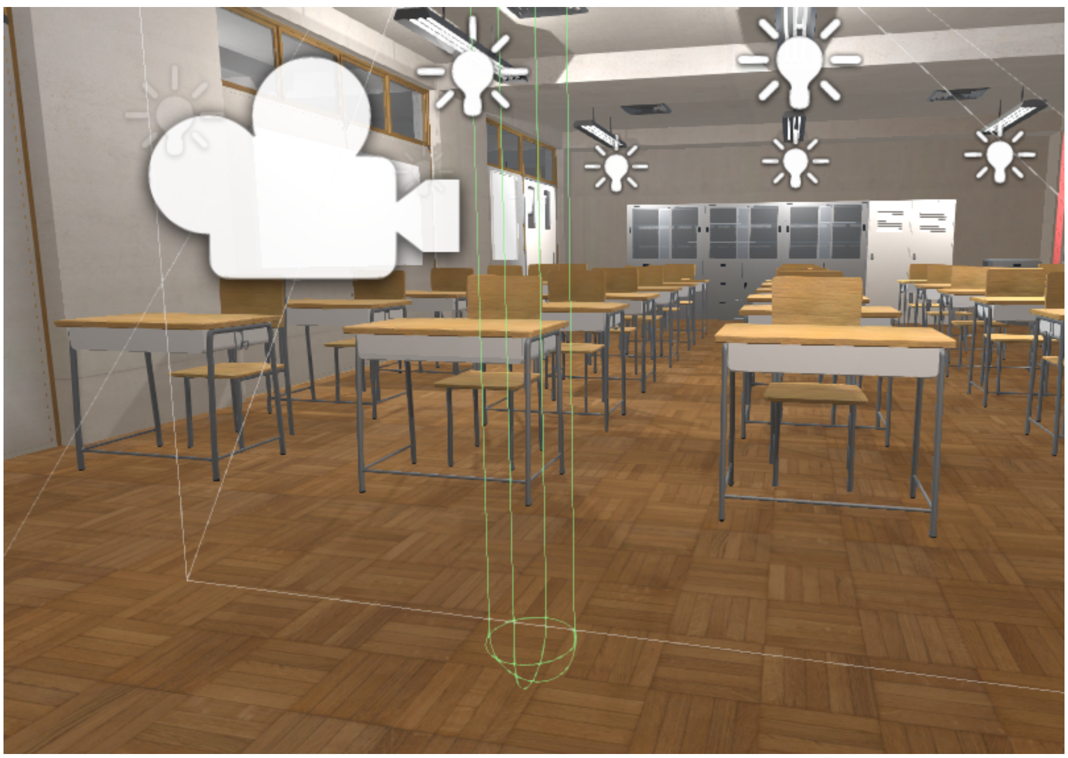

Figure 18: New player model collider in Unity

5.1.2 Changes to the execution flow

Because of the new attencion centered approach the project has taken, the existing flow of execution needed to be redefined. The multiple pauses and external, intrusive stimulus completely break the precious immersion we need to achieve, in order to obtain quality unbiased attention data.

Following this idea, we proceeded to eliminate the pause that took place in each scenario after the critical event had taken place. It did not make sense to eradicate from the experience the portion of time the participant spent elaborating a course of action, as it provides essential information on reaction time to solve a conflict, attention drift when the criticality occurs, movement (or lack of it) towards the problematic student that caused it, and future gaze fixations.

Along with the pause, the game used to show a list of possible actions the teacher could adopt in order to solve the conflictive situation, and a narration explaining the correctness of the action taken by the participant. Together with Barcelona's pedagogical team, we came to the conclusion that providing the user with a list of preset actions would potentially limit their freedom of thought, as the game would be actively coercing them to act in a certain way and manner.

Also, providing feedback at the end of each scenario could condition the way players behave in the next scenario, again deviating from the natural response that the participant would have (and in many cases discouraging them, as the most common feedback achieved had a negative connotation). For these reasons, both the auditory feedback and the suggestion list were removed.

With all these changes combined, we obtain a continuous execution flow with no stops, that also favors the participant's spontaneity and creativity when approaching critical situations, while at the same time paving the way to implement a 3D analysis software (Cognitive3D, explained ahead).

5.1.3 Immersion upgrades via audio

As stated multiple times, immersion is a crucial aspect of virtual reality experiences, and the power of sound in this matter is not to be underestimated. We used audio for two purposes:

Firstly, the game needs to be as intuitive as possible, and so, whenever the user is in need of specific information, it must be provided in the least intrusive way possible, so as to not perturb the rest of the experience. As surveys showed, a considerable number of users didn't quite understand when each scenario began, and where it ended.

To solve this, we implemented timed start and finish class bells, so that they would be sure to follow the flow of the program (that is to say, understanding when they should begin a lecture and when they could stop. It was frequent, especially in the second scenario, that participants asked the organizers what was going on, and obviously this totally broke immersion).

Secondly, it is important to always aim to reproduce the environment that is being portrayed in the virtual experience, this being a classroom, some ambiance noise had to be introduced. This mainly consists of light chatter at the beginning and at the end of the class, and after the criticality had taken place.

All audio samples and queues were edited with Audacity, an open source audio software that provides an impressive amount of features, of which we used some, like the high pass filter and the noise cleanse.

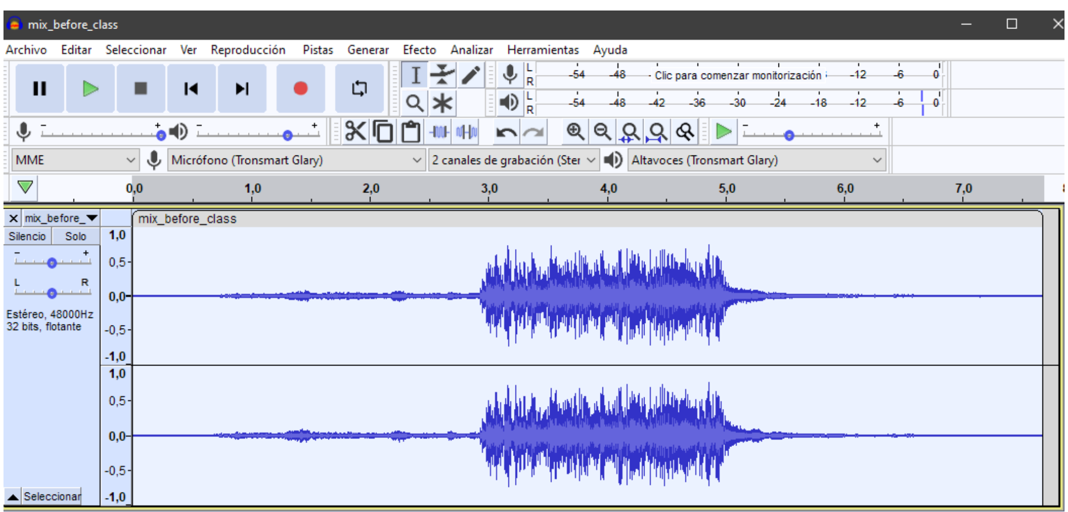

Figure 19: Audacity track after edits on sound queues.

5.1.4 Emotion analysis

Previously, the project counted with the voice analysis software known as Vokaturi. However, it was implemented several years ago, and its results did not provide any useful information, as it expected the voice input to be on english. Combined with the fact that the free version of vokaturi leaves a lot to be desired, it was decided that we would develop a new solution that gave accurate feedback when given spanish voice input.

The goal was to point out if a fragment of spoken dialogue contained either positive, negative or neutral tone, based on the contents of the speech. After exploring and considering different solutions, we decided to use two python libraries: Vadersentiment, and SpeechRecognition. Both libraries provide competent open source solutions to the problem at hand, being able to achieve the same functionality as the free version of Vokatury, but actually work with Spanish voice input, and because they are open source, long term support is almost warranteed, having a huge and thriving community. The code works as follows: First, an audio stream is opened as source and it's given to an instance of SpeechRecognition. After adjusting for ambient noise, it then starts recording for a given period of time. This recording is then sent as text to an instance of VaderSentiment (using function "polarity_scores' '), that assigns numerical values to the negative, positive, neutral and compound fields. The value of compound will determine the dominant field. If compound is greater or equal than 0'05, the dominant sentiment is positive. If it's smaller or equal to -0'05 it will be negative. The dominant sentiment will be set as neutral for the rest of the cases. Finally, the library matplotlib is used to create the pie chart.

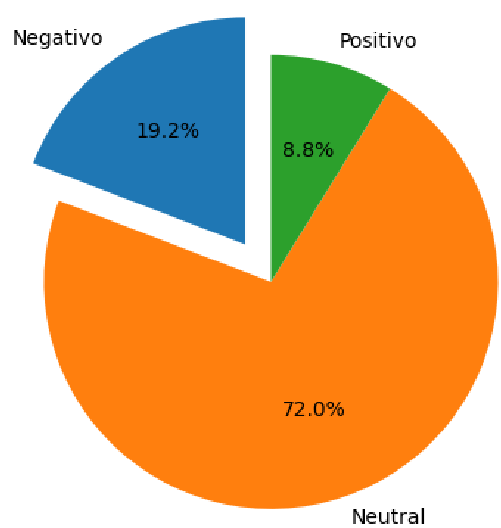

Figure 20: Output of the new emotion analysis python module

Although the code is fully functional, it is yet to be implemented along with the final solution of the Unity project. There are some constraints that difficult this process: Firstly, there are still some parts of the code that use Vokaturi's speech to text to determine when to fire action triggers, so there needs to be a replacement with the new solution (that is also able to translate text to speech), and secondly, the solutions cannot coexist with one another as for them to work, there cannot be two listeners at the same time (or else Unity launches an error). Due to other matters having greater priority, this is all the advancement made in this area (future work on this project shall consider finishing the implementation).

5.1.5 Pre-experience scene

It's important to take into consideration that a large number of users have not had a previous VR experience to our experimentation, and that in order to be able to gather real, unbiased data, there needs to be a scenario previous to the real experiment, to make sure everything is working as intended, and the user is confident around the digital environment we introduce them to. This scenario provides simple and concise explanations to the key aspects of the game, as well as tasks and feedback when completed correctly. The positive reinforcement provided when the task is completed successfully, further encourages the user to actively participate in the real scenarios, thus improving the data quality.

There had been multiple complaints on the user's end regarding the guidance when going through the experiments. Before, there was not a set up to properly introduce the participants to the game. Instead, they were always lectured from outside the virtual environment by the organizers, who would also answer questions related to the playability. Thus, the motivation for creating a guidance scenario.

The guide begins in a classroom with only one student. A narrator starts introducing the basics of the game, such as head motion to visualize the environment, player motion with the controller joystick, and player to world interactions. After a few moments, the player is asked to greet the student. When this is done, a completion sound plays and the final aspects about voice interaction and parameterization are explained. If the player does nothing, the program will stall indefinitely.

The interactivity of the guide was introduced so that when successfully completing the guide, the player would feel more confident, comfortable and safe around the game's environment, which is an essential requirement to ensure that the data collected resembles the real life behavior of the participant. At the moment, the guide can be found on the scenes folder, inside the unity project. It is fully operational although it does not belong to any menu yet.

Figure 21: Screenshot of the guide's inside perspective.

5.1.6 Analysis of 3D spaces - Cognitive3D

One of the most crucial parameters when trying to analyze body language is the behavior of human sight. Where we look, for how long, and with what frequency we fixate on an object, can determine a great deal of information about us. The chosen tool for carrying out the gathering and post-processing of data, is Cognitive 3D. An extremely competitive suite that enables us to add a new complexity layer to the existing project, focusing on the player's attention when going through the game.

Once selected as the main pillar of our parametrization journey, Cognitive3D had to be implemented on the already existing Unity project from Didascalias, focusing on the main feature that it provides: Gaze Tracking and Fixations. For this goal to be accomplished, some aspects of the morphology of the class needed to be adapted. These aspects could be grouped into freedom of movement (which was explained earlier on section 5.1.1), and gaze vector reachability.

Gaze vector reachability refers to the capacity of the user being able to look at something anywhere in the environment, and the corresponding gaze vector reaching the exact same point in space, so that it can be gathered and processed for the final data construct. Achieving this was considerably more complex than anticipated. In order for the vector to travel freely, we discovered that the student's invisible collider needed to be adjusted. This collider is used to interact with students through distance. It's an important aspect, so it couldn't simply be removed. Instead, it was reshaped and relocated so that a natural gaze would not intersect with it.

Figure 22: Comparison of the old collider (left) and the new collider (right).

For collecting gaze data, Cognitive3D is able to use multiple interfaces, of which we selected Fove, for its compatibility with the hardware we use on the experiments and its reliability. Fove uses an entrance vector to precisely determine what the user is looking at (this is why modifications needed to be carried out on the class environment: for Fove to work correctly). When any scenario is executed, the tool starts collecting data at a fixed rate (it can be adjusted, but performance needs to be taken into account), and when the scenario ends, the tool ends the session. This session is soon after uploaded to the Cognitive3D dashboard. Note that for the tool to work, it also needs to have the scene on which the data will be read, uploaded to the cloud service provided, together with the dynamic objects that will be present on the scene. Each dynamic object should have a unique identifier or else, somo objects migh appear as missing on the recreation (this means that all dynamic objects needed to be renamed appropiatly).

The Cognitive 3D dashboard provides access to the visualization of session data. This includes fixation history, gaze hitmap, movement recorded and gaze recordings. All of these data is represented within the same 3D scenario where the session took place (that is to say, the set of assets corresponding to the scene, is uploaded to the dashboard to provide an accurate recreation of the eventualities), and it can be visualiced as a 3D video, where one can move arround freely to inspect the scene. Other useful functionality this platform provides, is the capability to overlap multiple session recreations over a single 3D video, making it easy to compare the performance and/or behaviour of, for example, different users on the same scenario, or the same user through a set of diverse scenarios. Sessions may be grouped arround the versions of the 3D environmet they took place on, and also tagged and named dynamically.

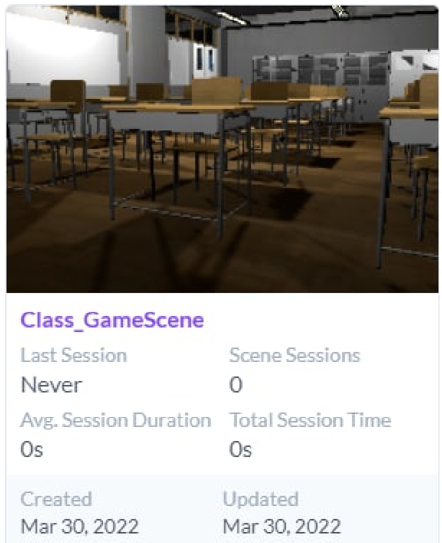

Figure 23: Loaded scene inside Cognitive3D's dashboard.

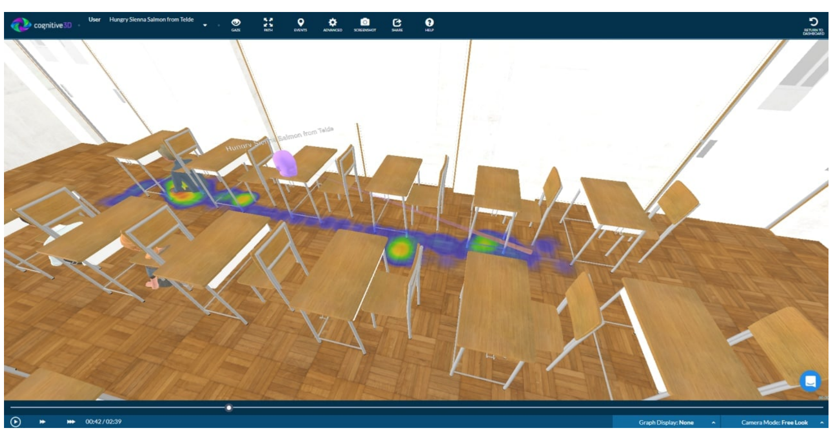

Figure 24: View of the 3D video recreation of the player's session.

In the near future, the project could take advantage of some other functionalities that Cognitive3D offers, specially the aspects regarding dynamic object interactions (that is, the students in the classroom or even new objects commonly found on class environments), and sensor integrations, like blood pressure and oxygen saturation, that would help further capitalize on the parametrization of the player's experience.

5.2 Performance upgrades

Performance is frequently overlooked when it comes to software development, as there are so many things to be accomplished in a limited time frame. However, when the complexity of a project increases, so do the hardware resource requirements, becoming more demanding for the system with each addition. This is particularly noticeable after implementing the 3D analysis tool. In this section we will describe the drive behind the performance improval, the tools used to determine the changes, the project modifications and the results this process has given

5.2.1 Importance of computer performance

As previously stated, the cumulative additions to the project had led to a high performance demand. A demand that could soon no longer be met due to the intrinsic limitations of a mobile set up (the experiments are always conducted on educational institutions, so only a portable set up makes sense). Performance issues began to get more noticeable after the implementation of the last features: Loading times between scenes were increasing, the frame rate was far away from reaching the standard 60 fps recommended for a smooth experience, and when making violent head movements, some participants started noticing black areas on the edges of the virtual reality headset's display (most likely due to the system not being able to render the complete size of the view in time, thus creating black zones on the headset's internal screen). These problems only got worse after repeated executions, as the system components' temperatures got higher, and thermal throttling could be noticed without any software profiling tool.

Of course, one can not expect a great immersion under these circumstances. And with poor immersion, the player is sure to diverge attention, go as fast as possible through the experience, or directly having to stop due to various discomforts. Improving performance is also beneficial to the projects' future implementations, as it leaves more room for integrating additional components before saturating the system.

5.2.2 Profiling tools and test bench

The system that would be used throughout all the performance evaluation, tests and Unity modeling describes as follows: The CPU is an Intel Core i5-6600k overclocked to operate up to 3.7GHz. The graphics card is an Asus Geforce GTX-970 (also overclocked), connected to an Asus Z170-A motherboard and counting with 16GB of DDR4 memory. In total, the system has 7 cooling fans, with 3 behind used on the CPU, 2 used for the GPU and the last 2 used for case airflow. It's a desktop configuration with excellent cooling, a fact to have in mind when comparing with laptop configurations that usually lack both the muscle and the cooling capabilities of a desktop PC.

For the software tools, we used Speccy to track temperature shifts, as it lists components individually (which helps when looking for bottlenecks) and Unity's built in profiler, as well as the statistics panel for ingame data. These two will provide data regarding frame rates and its dips and heights, as well as showing a representation of CPU and GPU usage accross the execution. Other software options were considered (such as Intel's Graphics Monitor) but due to results being clear enough with the aforementioned tools, extra tools were not needed.

5.2.3 Project modifications

We discovered that perhaps the most demanding aspects of the game performance wise, are the following: lights and shadows and the way they are casted, dynamic objects and general textures (this includes static objects and dynamic objects). And so we decided to focus on these aspects to improve the performance of the game while maintaining its appearance, or maybe even improving it.

- **Texture and shader adjustments**

  First, we addressed the textures. For each material, we changed the selected shader. The shader that was previously selected for all objects of the game, was the default one that Unity sets when importing an item. Our selection was the VertexLit shader for all materials, as it offers great performance improvements with no visible difference (this shader is very well optimized for mobile devices). Also, the reflections checkfox was activated, so we changed that too. Reflections add considerable amounts of rendering load to the system, an effort that should better be saved for the lighting (again, this was set as the default option by Unity). We also lowered the camera's clipping planes, as the environment in a relatively close distance, so there is no need to render everything from afar.

- **Processing objects in batching modes**

  Dynamic objects are a complex topic to address performance wise. Together with the texture adjustment, the other factor that we could modify is batching. Unity uses batching to group several objects to minimize the state changes needed to draw each object inside the batch. Doing this, in turn, leads to improved performance by reducing the CPU cost of rendering the objects [35]. There are two types of batching, depending on the nature of the object; Dynamic Batching and Static Batching, respectively for dynamic objects and static ones. Only objects that share common properties can be batched together. Static batching is the recommended technique for objects that do not move since it can quickly render them.But when it comes to dynamic objects, batching larger meshes dynamically is more expensive than not batching them, and so by deactivating dynamic batching, the CPU was freed from some degree of stress, depending on the number of students on screen.

- **Lighting and shadows**

  Last but not least, the changes of lights and shadows. Lighting is an extremely complex factor in game engines, as they shape the way everything looks. This, of course, comes with a price, being the most intensive task for the computer to handle. The first item to modify is directional light. We changed its mode to mixed instead of real time, as this would save extra calculations, and made its shadow to be only hard shadows, instead of the presented hard and soft shadows. We also lowered its intensity. The general lighting render of the project was also modified: Pixel Lights determine shading for each pixel of the object being shown, and pixel light count is the parameter that controls this value. It was originally set to 4 (scale goes from 0 to 4). We lowered the value to 3, seeing enough performance gain while maintaining the visual look of the scenario, together with the halving of the light resolution. Options for anisotropic textures, anti-aliasing and soft particles were also disabled.

- **The importance of baked lightmaps**

  Lightmapping is the process of pre-calculating the brightness of surfaces in a scene. It stores the information it calculates in a chart or lightmap for later use. Lightmaps allow adding global illumination, shadows, and other lighting at a relatively low computational cost [36]. To make full use of it, it was required to configure a new lightmap, with lower resolution, and lower size (as a small scenario does not need the default size), and also using the GPU to balance the load (Lightmapper can use GPU resources to free to some extent the CPU load). After selecting all the game's models to generate new Lightmap UVs, we could finally generate the new lighting (baked lightmaps UVs are used all across the Unity project to improve performance, except for dynamic objects). As a last touch, each of the 14 point lights had their shadow cast changed to hard shadows only (lowering as well the shadow resolution to half the original, and switching to distance mode, that can be adjusted to only be calculated for a set amount, instead of indefinitely, as it was), and their radius was adjusted for avoiding light overlapping (this implies that the intensity of each individual point light also had to be tweaked until the desired illumination was achieved). The shadow cascade remains the same, as modifying it resulted in major visual changes (it has 4 cascade splits).

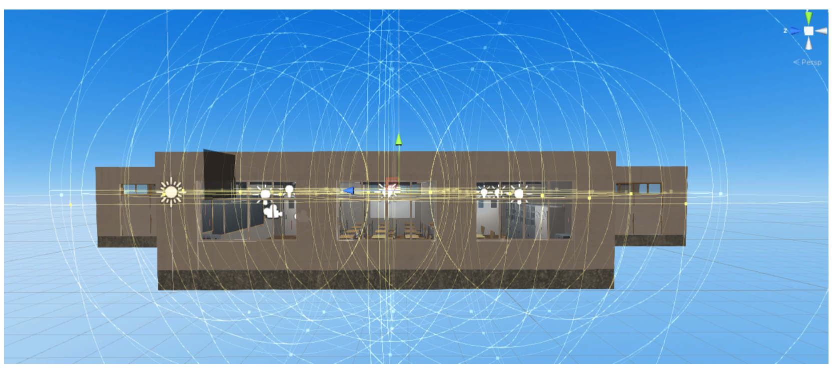

Figure 25: Outline of all point light area of effect around the environment.

5.2.4 Before and after comparison

The resulting gains in performance will be divided into frame count, and thermal readings, as these are the most important aspects for user experience and extender performance over several, continuous program executions.

As can be seeing on the figure, both the frames per second and the rendering times have vastly improved, the new version having 7 and a half times more frames per second than the original version at the most demanding event (that is, when the class is filled with the highest number of students), and the CPU rendering time taking around 6ms to render each frame, compared with the previous 50ms.

On the Unity profiler, notice how the dotted blue line does not get crossed on the profiled version of the project, whereas on the previous version, it is crossed and the game finds trouble staying well above the 60 frames per second mark (represented by the blue dotted line). It can also be seen that there are almost no performance dips on the improved version, staying at a flat count most of the time, unlike the older version that is constantly fluctuating back and forth (probably due to poor load balancing between the CPU and GPU).

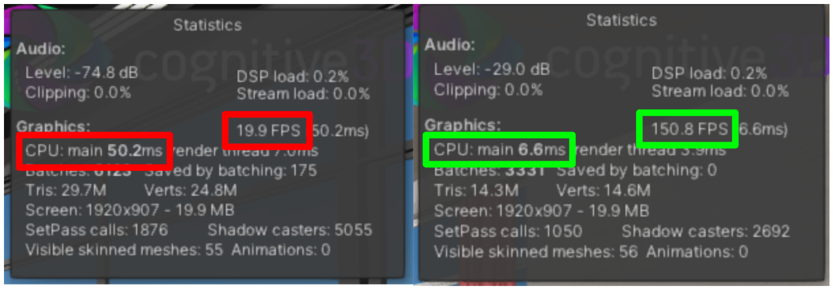

Figure 26: Performance comparison (FPS and CPU render time) on max load.

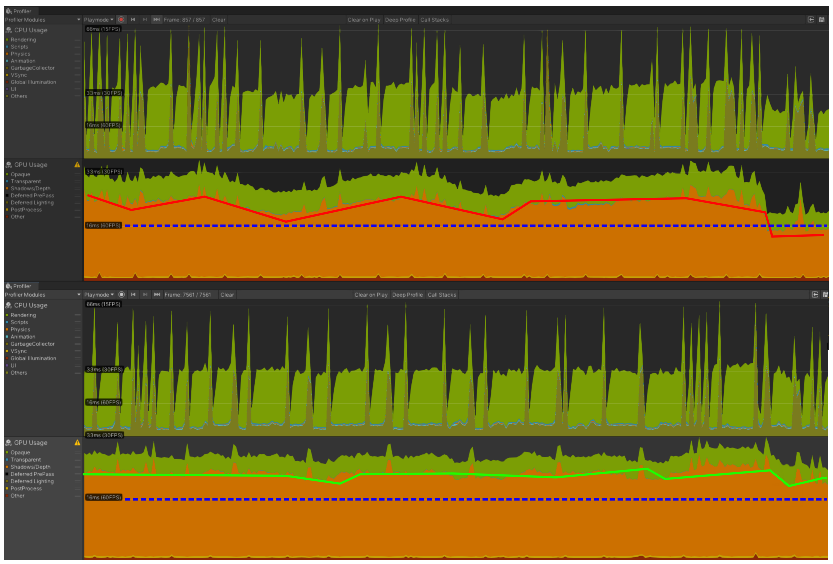

Figure 27: Organization of the Google Drive Unit of the project.

Significant improvements can also be noted on the thermal reading over several program executions with no idle time in between. It should be noted that when comparing a laptop with a desktop, we can expect more radical readings, as the cooling in portable computers is notoriously worse. The thermal improvements also mean that getting the system to a thermal throttling state will prove to be far more difficult, concluding a better performance when sustaining several runs in a go.

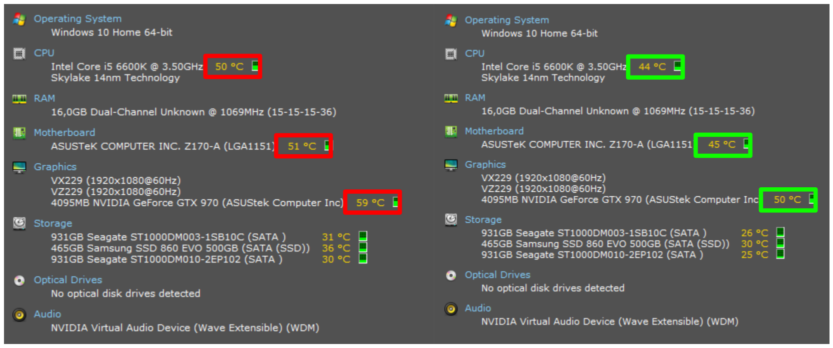

Figure 28: Comparison of computer temperatures after several program executions.

6 Experimentation phase

In this last phase, we decided to carry out a new experiment independent of the two that had been done previously, following a similar but improved structure, based on the learning we had acquired from the previous ones. In this case, the importance of the experiment was to test the incorporation of the new tool that we had introduced to the project. Therefore, we devised clear objectives to take into account during the experiment, so that we would focus on the new data to be analyzed, and not lose sight of this last test purpose.

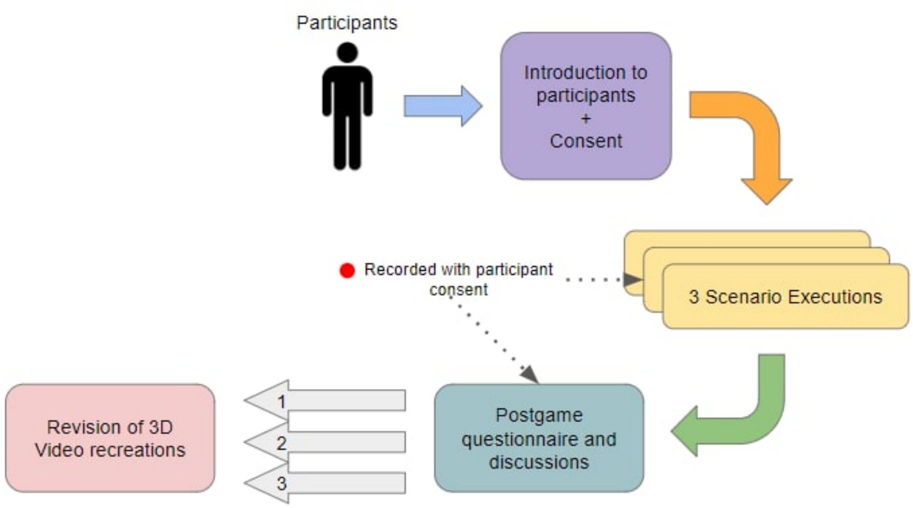

Figure 29: Experimental design diagram.

6.1 Objective definition

Once we successfully integrated Cognitive3D into the project, and after making all the necessary changes to the application, we moved on to the experimentation phase. On April 7th we went to the Ramón y Cajal School to be able to test the new changes.

Our main objectives were the following:

- **Full tool operation**

We had to check that the ClassroomVR app was working perfectly as many changes had been made. Not only in terms of technical operation, but also as a learning tool. The experiment together with the participant surveys would give us interesting feedback regarding its proper functioning.

- **Cognitive3D operation and utility**

A clear and important objective was to check the usefulness of the new tool that we had recently included in the project, Cognitive3D. Apart from verifying its correct operation and that it did not interfere with any of the other functionalities, it was important to verify that the data obtained was correct and also useful for our investigation.

- **Compare and contrast data**

We had to contrast the data obtained in this new phase with the previous ones, in order to carry out an analysis based on the research carried out throughout the project.

6.2 Comparison with previous experiences

Unlike previous experiences, one of the most noteworthy points was that this time we did not have the physical support of the Barcelona team. Therefore, we had to fully develop the experiment with our skills and experience. This was not a problem, since the two previous experiences had helped us to fully understand how the experiments worked. In addition, we had already gone through all kinds of cases in the previous ones, thus having a wide range of solutions in case any problem arose in the middle of the experiment.

During this experience, there were no notable problems. The experiment was carried out satisfactorily, and all the teachers who participated voluntarily left the test with a good feeling. In addition, it should be noted that many of them were very interested in the project and how it would continue in the future.

6.3 Experiment in Madrid (II)

Our main goal for this last experiment was to test the usefulness of the Cognitive3D implementation we had just made to the project. As an important piece of information, we designed a structure around this new implementation, conducting an interview at the end of the test in order to compare the data obtained.

During the preparation of the experiment, we spoke with the director and the coordinator of the educational center, who helped us to choose the best space to carry out the test. This space was a room protected by soundproof glass belonging to the library, giving us a space totally free of external noise that could be a distraction in the experiment.

Once the place of development was chosen, we began with the provision of the equipment that we were going to use. We decided to free the space of all possible furniture that would hinder the participant during the execution of the experiment.

After the placement, we carried out a complete test on our own to verify that the project worked perfectly, and thus avoid possible problems with the first participant. When we approved all the material and the preparation, we decided what role we would take during the entire experiment, dividing the tasks in the most efficient way.

The study involved 9 teachers of Ramon y Cajal school in Madrid. All the teachers participated voluntarily; this way all the tests started from users with interest and proactivity in the subject. Within this group of people, we find both young and less experienced teachers, as well as veteran teachers who have more background. Two thirds of the participants were teachers with 10 or more years of experience in the field, with the remaining third of participants having less than 10 years of experience. Both cases are important for the experiment and helped to draw different conclusions. Regarding gender, 4 participants identified as females, and the other 5 identified as males, although this characteristic did not imply any significance across the experiment.

The experiment was carried out in the following way:

- **1. Test preparation and execution.** First, the participants were told the basics about the experience. Afterwards, they were moved individually to an approximate area of 2 square meters, to ensure freedom of movement, and they were fitted with the virtual reality headset (previously disinfected, as per covid regulations, as can be seen on Figure 1). Before the program started running, we started screen capture to have testimonial footage of each participant's run (with their consent, of course).

- **2. Questionnaire.** After the participants had gone through all three different scenarios the game provides, they were asked to take a quick survey, and the conversation with the corresponding interviewer was also recorded (again, with the participant's consent). The survey consisted of questions about the general quality of the experience, along with questions about their previous experience with virtual reality, and most importantly, a question relative to their gaze awareness during the test.

- **3. Final step.** As a final step, the participants were presented with their 3D recreation of gaze and fixations (that is, the final 3D construction representing all the gathered data by Cognitve3D), with the intention of finding out if their actual focus matched with the feeling of awareness each participant had during the VR experience.

As for the instruments regarding the experiment, we used an MSI laptop with a dedicated graphics card, a pair of Oculus Rift S virtual reality headset, and a portable recorder. This is all meant to facilitate mobility, as all experiments are to be done on educational entities. On the software side of things, we used the previously mentioned Cognitive3D and its online dashboard, along with Fove interface (which is integrated as a Unity module), Windows' 10 native screen capture utility, Google Forms for the questionnaires, and of course, the Didascalias Unity project.

Figure 30: Teachers completing the 3D experience (experiment II Madrid).

6.4 Second experiment conclusions

- **Data collected using Cognitive3D:**
  - **Motion gaze control:** Data we extract through the tool that shows where the user's gaze is at all times. Thanks to this data, we can compare the different tests of each of the participants within Cognitive, emphasizing the different places and students on which they have chosen to focus their attention.
  - **Heatmap:** A heatmap that shows a general oath of gaze focus during the test. It allows us to know where the gaze has been fixed the longest, by marking it with reddish tones, as shown on Figure 2.

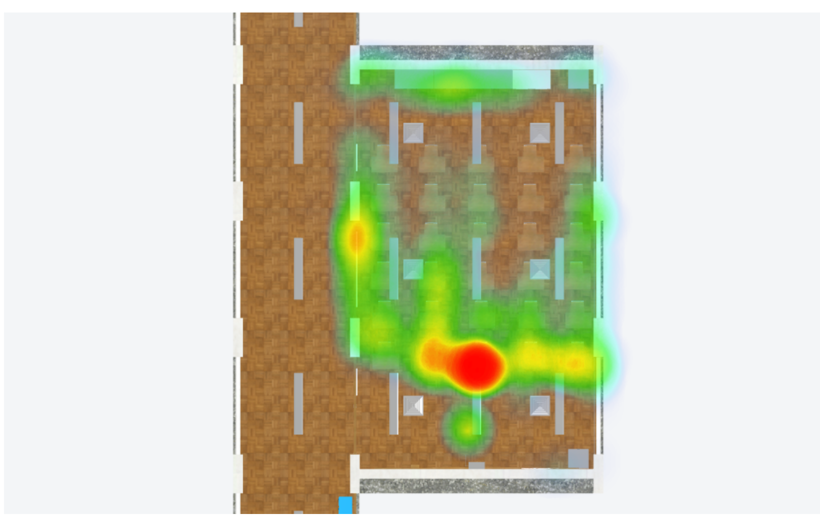

Figure 31: Zenith heatmap of aggregated gaze data.

– **Cube collision map:** We also extract a map that allows us to know at what points his gaze collides. As can be seen in the image, it is shown on Figure 3 that he has focused his gaze on the student at the second desk, along with the one on the far right.

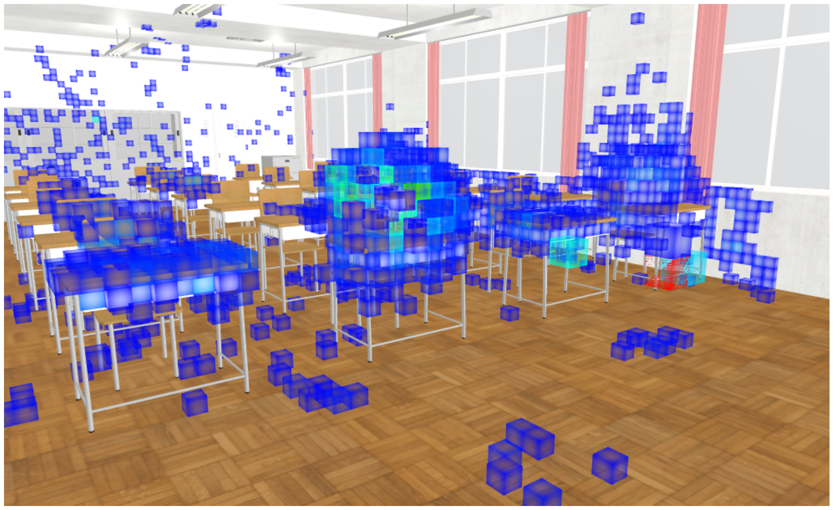

Figure 32: Cube collision map representation.

– **User path:** We take in count the user's movement path, allowing us to see which users have moved the most and where to, as can be seen on Figure 4. In this way, we can see how it affects their attention in the class if they choose to move a lot or to stand still at their desk.

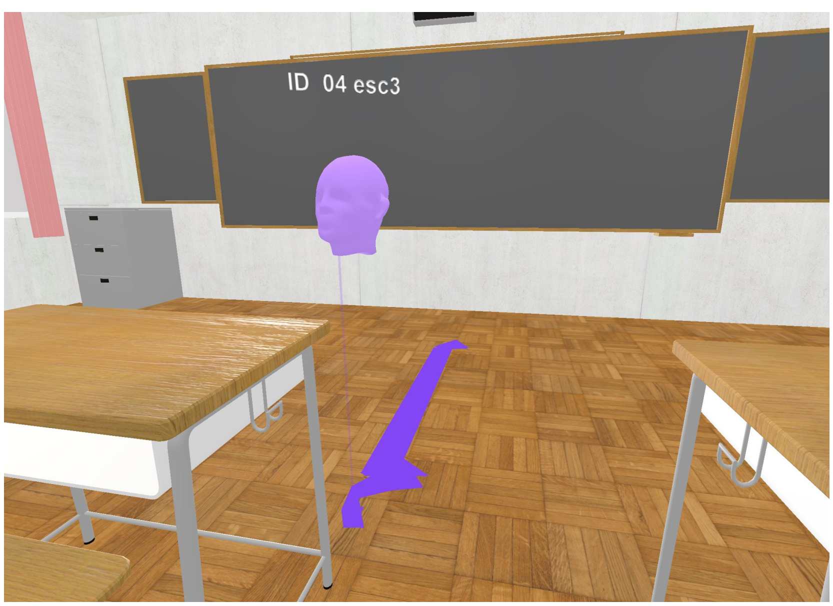

Figure 33: Player movement representation.

Question + Rating (1 5)	1	2	3	4	5
During the experience, I was aware of where I was looking at 100% of the time: (1- Not conscious at all, 5-Aware at all times)	0%	11,11%	11,11%	11,1%	66,67%
I consider the quality of the immersion to be: (1-Not immersive at all, 5-Almost like real life)	0%	0%	33,3%	55,6%	11,1%

Table 2: Questionnaire results

- **Session comparison:** We also compare how users behave in the same scenarios, representing their positions and their gazes at the same time. This way, we can directly see the differences in how each one has faced the stage, and how those decisions affect the attention of their gaze.

- **Questionnaire statistical data:** A simple questionnaire where participants were asked, among other things, to which extent were they conscious about their gaze positioning during the game. Here are the most relevant question results for the scope of this paper:

- **User feedback for future work:**

During the final interview, the participants gave us very enriching feedback to take into account for future iterations of the project.

For the most part, they understood the scenarios and the conflicts without problems. However, most of them highlighted the emphasis on creating new scenarios with another conflict perspective, in order to have more specific cases. They pointed out that when a student rebels in class it is very common for the other students to follow him, or when a conflictive situation occurs on the part of someone, the other students tend to disperse even if they do not follow the game. Pedro Uruñuela, expert in coexistence and author of [The Classroom Management], addresses the phenomenon of school disruption, affirming that the disruption and even expulsion of a student is very negative for the whole class in general. Therefore, it is true that when resolving a conflict, it is not only necessary to take into account the students involved in it, but it is important to attend to the whole class.

Most of them also commented that dark bands were sometimes seen around their vision in the glasses. This phenomenon is due to the fact that it takes time to process the data of the scenario and the execution, and therefore those dark bands are seen while loading. Due to this problem, which we had already encountered the first time we did the project, we decided to give importance to improving the performance of the project. As can be seen in the performance section, we achieved a huge improvement from 19.9FPS (50.2ms) to 150FPS(6.6).

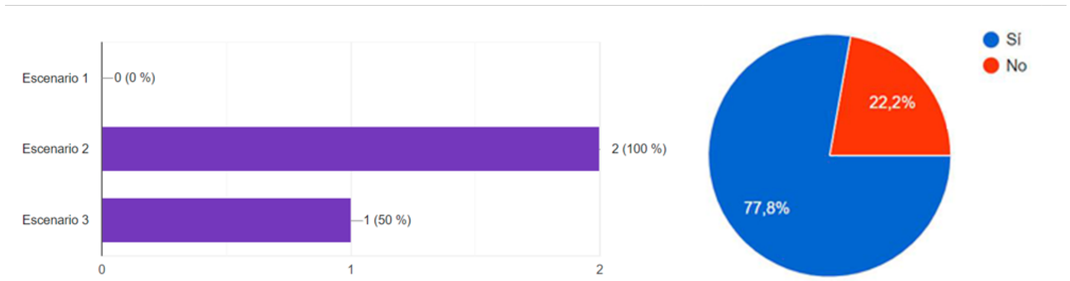

Figure 34: Feedback data of the participants on the understanding of the scenarios and their conflict.

7 Conclusions and future view of the project

7.1 Conclusions and future work

ClassroomVR is the tool in which we have entered our efforts during this Final Degree Project. Its main objective is to improve communicative aspects, both verbal and non-verbal, to help them more efficiently manage the conflicts that can arise in a classroom. For this tool we have made changes proposed by the Didascalias team, improvements proposed by us as a team, and we have also provided a new approach to the analysis of teachers through the Cognitive3D tool.

At the beginning of the project, during the experimentation phase, we focused on understanding the tool well. Thanks to the two experiments carried out in this phase, we were able to verify that the tool is really very useful when it comes to presenting conflicting cases that cannot normally be represented as a teacher's learning exercise. This is due to the fact that both the learning professionals in the first experiment, as well as the students of the Teaching Master's Degree in the second, were totally gratified by the experience, as shown in the interviews carried out.

However, the final conclusions of our research and contribution to the project came after the experimentation phase, once we were able to start collecting data on the attention of the participants within the test.

The first thing that struck us was that most teachers keep their attention only on the front row of the virtual class, as can be seen on the heatmap, even when the problematic scenario is caused by a student at the back of the class. This might be because we humans find an easier time establishing communication channels with people physically closer to the emitter [37], (although this is yet to be confirmed in further evaluation). We could assume, seeing as the quality of the immersion was positively rated by the participants, that this behavior is to be extrapolated to a real-life classroom environment, possibly implying that students may be highly aware of the lecturer's attention dips and points of major focus.

We also found another great point to take into account, as most participants stated that they were attentive to the conflict of the scenario, but data shows that the situation leads them to disperse their gaze in different ways. With this data we can point out that, in this type of situation, people tend to carry out small actions or gestures of which they are not aware. As we have explained before, these actions cause their gaze and attention to decrease. But according to the participant's interviews, when the students notice changes in the teacher's behavior, they tend to disperse even more. Thus, there is indirect feedback generated that may harm both parties.

Taking this into account, when comparing the data from the interviews with the data from Cognitive3D, it is shown that the more experience the teacher has, the less their gaze is diverted at these critical situations. Therefore, we conclude that the teachers who have faced this type of situation the most, see their attention less disturbed, and so, this kind of simulation has a positive effect on lecturers when they are preparing to handle the classroom, but don't have access to this kind of experience.

As for the future of the project, although the achieved parametrization of nonverbal language is limited, this project has not only provided us with useful data on the subject at hand, but has also given us different ideas to consider.

One of those ideas, if not the most interesting field to have in mind, is the parametrization of facial expressions, this of course being a tremendous source of information regarding human body language. This new aspect promises to compensate for the limitations of gaze-based assumptions and suggests new possibilities of study.

Finally, based on the feedback from the participants during the interviews we have considered the possibility of implementing new scenarios to further capitalize on the variety of hassles that may occur in an educational setting. These proposals will be evaluated by the Barcelona pedagogical team, who will be able to provide greater insight into the situations and their possible benefits.

As a final touch, we gathered all the knowledge and procedures in a paper that would be published on the 30th International Conference on Computers in Educations (https://icce2022.apsce.net/). Here, we expose our work regarding attention based on gaze movements, and it's potential impact over an educational environment, along with some ideas for future work regarding studies based on tracking the teacher's non-verbal language parameters (that is to say: body posture, facial expressions, vital signs such as blood pressure and heart rate, etc).

8 Annex

8.1 Submission of a paper in the ICCE 2022

In order to publish the research carried out in this work, we decided to write a paper which we named *Study on attention and interaction within computer generated 3D educational environments*, which we have sent to the 30th International Conference on Computers in Educations (https://icce2022 .apsce.net/).

The 30th International Conference on Computers in Education is a conference organized by the Asia-Pacific Society for Computers in Education (APSCE). The theme of this conference is Optimizing technology for sustainable quality education in the new norm, so we thought that the research of our project would be interesting within that field.

The article collects the most relevant sections of the project, with special emphasis on the experimental part, where we carry out our research based on gaze movements, and it's potential impact over an educational environment. Using the data collected up to that moment, we were able to carry out the research, drafting a structure that touches on all the important points to perfectly understand the investigation carried out and the reason for it. Thus, we were able to reach conclusions based on the specific points of the article.

Study on attention and interaction within computer generated 3D educational environments

Bruno MAYO[a], Fernando MUÑOZ[a], Guillermo MONSERRATE[a], Alejandro ROMERO[a], Ibis ALVAREZ[b], Borja MANERO[a]
[a]*Universidad Complutense de Madrid, Spain*
[b]*Universidad Autónoma de Barcelona, Spain*

Abstract: Currently, despite the importance given to teacher training, there are not many tools or activities that help them practice in conflict situations that may arise in any classroom. It has been proved that one of the most important problems generated by these situations is the general loss of attention on the part of the students when seeing an inexperienced teacher react to the conflicts. This paper presents the analysis of a virtual reality environment designed to improve the communication skills in secondary school teachers, putting special emphasis on the participant's attention during the simulation in order to help them not to lose control on their students.
The system captures the user's gaze, showing all the points they have paid attention to, and for how long. This way, at the end of the execution, the user can be aware of where their attention was focused in each case. To test the effectiveness of the tool, we carried out a study through 9 educational professionals. The research allowed us to prove that most of the participants had not been aware of where their attention was during the test. In addition, thanks to the results of the professionals that had more experience, a better control of their attention was noticed, proving that the experience in this type of cases has helped them to acquire a better manage of the conflictive situation. Thanks to this study, and the opinion of the professionals that we consulted, we are positive that this tool can really help future teachers to improve their abilities in their educational formation.

Keywords: Education, non-verbal communication, virtual reality, classroom environment, teacher formation, serious games.

1. Introduction

Virtual reality is not considered a new technology anymore. In parallel, Its use in education has been present since the middle 60's where it was recorded to have been used as part of the training program of air force pilots (Kavanagh et al., 2017). Although this technology used to be unavailable for the public sector, nowadays new Head Mounted Displays (HMD) such as Oculus Rift and HTC Vive have been developed and put out in the market at reasonable prices making the use of these devices open for the masses.

All these factors have opened the doors to a new world of possibilities where new studies are being conducted (Hedberg & Alexander, 1994) since the entry price for them to be carried out is lower. Although this has been theorized for some decades, now is the best time to get to the practical side of the matter and get tangible results of the previous written work.

On a different vein, numerous articles and studies have been published relying on body language parametrization and analysis to answer questions not necessarily related to virtual reality itself but also about social aspects of humans, the use of these devices in medicine and many other subjects (Ahir et al., 2020).

Some of these studies talk about the use of certain tools (Metallinou et al., 2016) some others talk about interactions between humans (Volkova et al., 2012) and even present in the medicine field (Kaltenborn & Rienhoff, 1993). Since the scope of the use of these tools is so broad, we saw a good opportunity to elaborate our study with the help of some of these tools. This will allow us to get more visual results, making it easier to elaborate on the result of the experiments later on.

The use of this hardware and software, such as motion capture suits and word recognition applications (Fuertes Franco et al., 2020) , has proven to be very helpful and has made it easier for researchers to bring quantitative results to the table and analyze data otherwise very complex to determine.

On the other hand, research shows that in the present, teachers don't receive good quality training while they are preparing to give real lessons (Sáez-López et al., 2020). People enrolled in the primary and secondary education degrees in different universities state that the teaching material available for them lacks some groundwork (Cárdenas et al., 2021). Even though they have some practical training, they think it would be better if they had some real world experience with kids and teenagers in real classrooms. Since during practical sessions they have the support of an experienced teacher, they claim the situation would be much different if they had to face some of the conflicts and other issues that may occur during the lessons.

Some studies show the trend of using games for different purposes other than entertainment is going up year by year since various decades ago (Susi et al., 2007; Zhonggen, 2019).These serious games refer to games (whether is board games, videogames or any other kind) developed for an educational purpose where the players' ability will unlock new elements and teach them subconsciously a new skill or measure their capabilities at some tasks. For example, Didascalias project developed a serious game for future teachers to build their confidence and skills solving conflicts in the classroom. By putting the subjects into a close-to-real-life situation based on pedagogical studies, both students and researchers can see how they would perform in a real classroom where, for example, a conflictive student refuses to join a work group. We thought this would be a very good starting point to develop our study since body language takes a big part in these situations where the teacher needs to keep his/her authority for the lesson to run smoothly.

The main goal of this paper is to provide an insight into the interaction of people with the 3D environment and where their focus is aimed at during the different situations. In our case, teachers who interact with students in the classroom and the conflicts that may take place in it. In order to accomplish this, we carried out an experiment involving secondary school teachers, who will be our test subjects, and using data-collection applications which we added into an existing project. The results will give us a clear view on how people move, direct their sight and how they pay attention to their surroundings among other interesting data useful for our project.

This document will be structured as follows: Section 2 will overview the adaptations introduced to Didascalias project, then, Section 3 will show the experimentation phase where the key elements of our objectives will be put to test, Section 4, will review the results of these experiments, and finally, some overall conclusions and some ideas for future work on the project. Additionally, a reference section will be added for convenience.

2. Project adaptation and evolution

As mentioned before, we started from an already existing project: Didascalias (Cárdenas et al., 2021). A project with the purpose of introducing new teachers to complicated situations that may arise in the class. We considered this would be the perfect ground to add a new layer of complexity. What if we could capture, process and analyze the crucial information we all give away by means of gaze and fixations? This premise is where 3D software analysis tools come in. First, we had to select the solution that best fits the educational nature of the project, within a very competitive and new market that has rapidly evolved to become the future of human - computer interaction: Virtual Reality. We used Cognitive 3D software to ease the parametrization process.

One of the most crucial parameters when trying to analyze human body language is the behavior sight (Edinger & Patterson, 1983; Sanders & Wiseman, 1990; Smidekova et al., 2020). Where we look, for how long, and with what frequency we fixate on an object, are dead giveaways for information, including of course, data regarding our attention. For this reason, the project needed to undergo modifications to ensure that users could freely look around the environment, and that this precise information could appropriately reach the tool in order to construct the final recreation.

To achieve this goal, it was necessary to rearrange the colliders used as action triggers for the program logic. These colliders, while both invisible and imperceptible by the user, interrupted the gaze vector (meaning it couldn't be represented as intended).

Another relevant aspect when evaluating a subject's attention on the upcoming experiments is providing sufficient interaction so that attention is focused inside the 3D educational environment instead of the outside world (real world) (El-Nasr & Yan, 2006; Kim & Biocca, 2018; Pallavicini & Pepe, 2019; Pfeiffer, 2012; Pfeiffer et al., 2008). This is mainly obtained by means of sound queues and freedom of movement, which, to be achieved, some important adjustments were taken into consideration.

As a summary: the teacher's movement was smoothened to minimize motion discomforts, and obstacles around the classroom, as well as the player model, got their colliders adjusted, keeping to a minimum possible frustration. These changes were carried out in an effort to further encourage participants to freely interact with the educational environment as they would in real life. Figure 1 illustrates the information flow of the program.

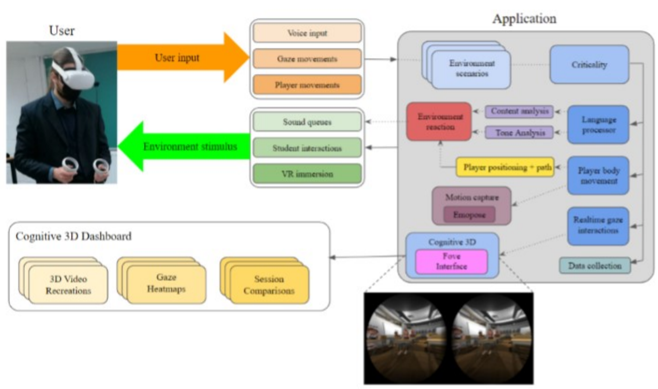

Figure 1. Program's information flow graphic description.

3. Experimentation phase

3.1. Participants

The study involved 9 teachers of Ramon y Cajal school in Madrid. All the teachers participated voluntarily; this way all the tests started from users with interest and proactivity in the subject. Within this group of people, we find both young and less experienced teachers, as well as veteran teachers who have more background. Two thirds of the participants were teachers with 10 or more years of experience in the field, with the remaining third of participants having less than 10 years of experience. Both cases are important for the experiment and helped to draw different conclusions. Regarding gender, 4 participants identified as females, and the other 5 identified as males, although this characteristic did not imply any significance across the experiment.

3.2. Experimental design

The experiment was carried out in the following way:

1. *Test preparation and execution*. First, the participants were told the basics about the experience. Afterwards, they were moved individually to an approximate area of 2 square meters, to ensure freedom of movement, and they were fitted with the virtual reality headset (previously disinfected, as per covid regulations, as can be seen on Figure 2). Before the program started running, we started screen capture to have testimonial footage of each participant's run (with their consent, of course).
2. *Questionnaire*. After the participants had gone through all three different scenarios the game provides, they were asked to take a quick survey, and the conversation with the corresponding interviewer was also recorded (again, with the participant's consent). The survey consisted of questions about the general quality of the experience, along with questions about their previous experience with virtual reality, and most importantly, a question relative to their gaze awareness during the test.

Figure 2. Teachers completing the 3D experience.

3. *Final step.* As a final step, the participants were presented with their 3D recreation of gaze and fixations (that is, the final 3D construction representing all the gathered data by Cognitve3D), with the intention of finding out if their actual focus matched with the feeling of awareness each participant had during the VR experience.

3.3. Instruments

As for the instruments regarding the experiment, we used an MSI laptop with a dedicated graphics card, a pair of Oculus Rift S virtual reality headset, and a portable recorder. This is all meant to facilitate mobility, as all experiments are to be done on educational entities. On the software side of things, we used the previously mentioned Cognitive3D and its online dashboard, along with Fove interface (which is integrated as a Unity module), Windows' 10 native screen capture utility, Google Forms for the questionnaires, and of course, the Didascalias Unity project.

4. Results

According to the data obtained, we considered it important to divide them in two groups: quantitative and qualitative.

4.1. Quantitative

- **3D Recreation**:
  - *Motion gaze control*. Data we extract through the tool that shows where the user's gaze is at all times.
  - *Heatmap*. A heatmap that shows a general oath of gaze focus during the test. It allows us to know where the gaze has been fixed the longest, by marking it with reddish tones, as shown on Figure 3.

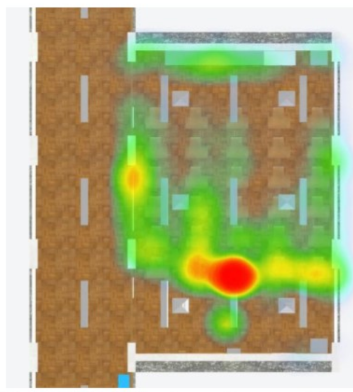

Figure 3. Zenith heatmap of aggregated gaze data.

  - *Cube collision map*. We also extract a map that allows us to know at what points his gaze collides. As can be seen in the image, it is shown on Figure 4 that he has focused his gaze on the student at the second desk, along with the one on the far right.

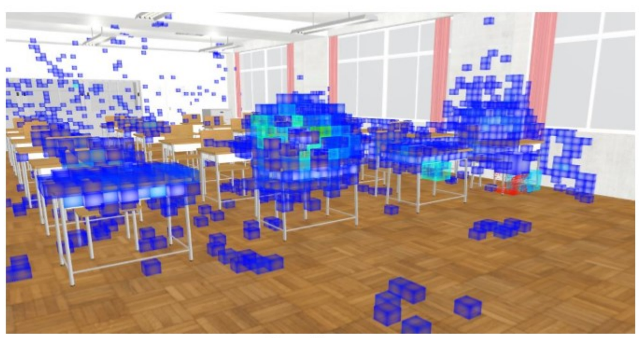

Figure 4. Cube collision map representation.

- *User path*. We take in count the user's movement path, allowing us to see which users have moved the most and where to, as can be seen on Figure 5. In this way, we can see how it affects their attention in the class if they choose to move a lot or to stand still at their desk.

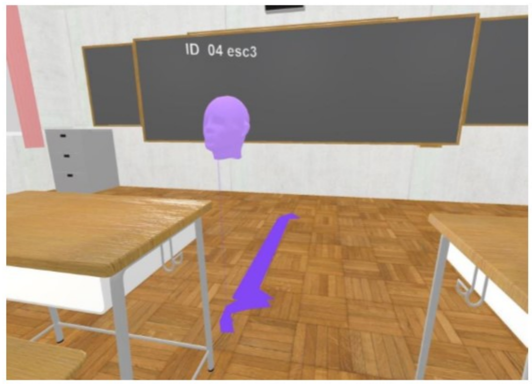

Figure 5. Player movement representation.

- *Session comparison*. We also compare how users behave in the same scenarios, representing their positions and their gazes at the same time. This way, we can directly see the differences in how each one has faced the stage, and how those decisions affect the attention of their gaze.

- **Questionnaire statistical data:** A simple questionnaire where participants were asked, among other things, to which extent were they conscious about their gaze positioning during the game. Here are the most relevant question results for the scope of this paper:

Table 1: *Questionnaire results.*

Question + Rating (1~5)	1	2	3	4	5
During the experience, I was aware of where I was looking at 100% of the time: (1- Not conscious at all, 5-Aware at all times)	0%	11,1%	11,1%	11,1%	66,7%
I consider the quality of the immersion to be: (1-Not immersive at all, 5-Almost like real life)	0%	0%	33,3%	55,6%	11,1%

4.2. Qualitative

This section of the collected data comes from the interviews conducted about the experiment and the participants' reactions to the Cognitive3D data. Therefore, we divide it into three points:

- ***General quality of the experience***. To prove that the test was optimal, and that the data can be taken into account.
- ***Previous experience with virtual reality***. To keep in mind their experience with this type of technology when analyzing the data.
- ***Gaze awareness during the test***. To analyze the most important comparative of the experiment. In this step, we show the data extracted with Cognitive 3D to the participant, and we ask them about it. This way, we can compare how their sensations have been with respect to the reality that the tool shows.

5. Conclusions and Future Work

This paper presents the integration of Cognitive3D in ClassRoomVR as a tool for analyzing and studying the attention of teachers within computer generated educational environments. The integration of this tool has been followed by an experiment to evaluate its potential when lecturers are exposed to conflictive, stressful situations.

The first thing that strikes us is that most teachers keep their attention only on the front row of the virtual class, as can be seen on the heatmap, even when the problematic scenario is caused by a student at the back of the class. This might be because we humans find an easier time establishing communication channels with people physically closer to the emitter (Kiesler & Cummings, 2002; Latané et al., 1995; Sanders & Wiseman, 1990), (although this is yet to be confirmed in further evaluation). We could assume, seeing as the quality of the immersion was positively rated by the participants, that this behavior is to be extrapolated to a real-life classroom environment, possibly implying that students may be highly aware of the lecturer's attention dips and points of major focus.

Most participants stated that they were attentive to the conflict of the scenario, but data shows that the situation leads them to disperse their gaze in different ways. With this data we can point out that, in this type of situation, people tend to carry out small actions or gestures of which they are not aware. As we have explained before, these actions cause their gaze and attention to decrease. But according to the participant's interviews, when the students notice changes in the teacher's behavior, they tend to disperse even more. Thus, there is indirect feedback generated that may harm both parties. Taking this into account, when comparing the data from the interviews with the data from Cognitive3D, it is shown that the more experience the teacher has, the less their gaze is diverted at these critical situations. Therefore, we conclude that the teachers who have faced this type of situation the most, see their attention less disturbed, and so, this kind of simulation has a positive effect on lecturers when they are preparing to handle the classroom, but don't have access to this kind of experience.

Although the achieved parametrization of nonverbal language is limited, this experiment has not only provided us with useful data on the subject at hand, but has also given us different ideas to consider in the future. One of, if not the most interesting field to have in mind, is the parametrization of facial expressions, this of course being a tremendous source of information regarding human body language. This new aspect promises to compensate the limitations of gaze-based assumptions and suggests new possibilities of study. Finally, based on the feedback from the participants during the interviews we have considered the possibility of implementing new scenarios to further capitalize on the variety of hassles that may occur on an educational setting.

6. Acknowledgements

This project has been funded by the Ministry of Science, Innovation and Universities of Spain (Didascalias, RTI2018-096401-A-I00).

7. References

Ahir, K., Govani, K., Gajera, R., & Shah, M. (2020). Application on virtual reality for enhanced education learning, military training and sports. *Augmented Human Research, 5(1), 1–9.*

Cárdenas, M. M., Álvarez, I. M., Romero, A., & Manero, B. (2021*)*. A Teacher Training Proposal for Classroom Conflict Management through Virtual Reality. *2021 International Conference on Advanced Learning Technologies (ICALT),* 373–375.

Edinger, J. A., & Patterson, M. L. (1983). *Nonverbal involvement and social control. Psychological Bulletin, 93(1), 30.*

El-Nasr, M. S., & Yan, S. (2006). *Visual attention in 3D video games. Proceedings of the 2006 ACM SIGCHI International Conference on Advances in Computer Entertainment Technology, 22-es.*

Fuertes Franco, E., Navarro Vaquero, J., & Montes Larrabaster, J. (2020). *VRetorik: un videojuego en realidad virtual para mejorar las habilidades de hablar en público.*

Hedberg, J., & Alexander, S. (1994). Virtual reality in education: Defining researchable issues. *Educational Media International, 31(4), 214–220.*

Kaltenborn, K.-F., & Rienhoff, O. (1993). Virtual reality in medicine. *Methods of Information in Medicine, 32(05), 407–417.*

Kavanagh, S., Luxton-Reilly, A., Wuensche, B., & Plimmer, B. (2017). A systematic review of virtual reality in education. *Themes in Science and Technology Education, 10(2), 85–119.*

Kiesler, S., & Cummings, J. N. (2002). *What do we know about proximity and distance in work groups? A legacy of research. Distributed Work, 1, 57–80.*

Kim, G., & Biocca, F. (2018). Immersion in virtual reality can increase exercise motivation and physical performance. *International Conference on Virtual, Augmented and Mixed Reality, 94–102.*

Latané, B., Liu, J. H., Nowak, A., Bonevento, M., & Zheng, L. (1995). Distance matters: Physical space and social impact. *Personality and Social Psychology Bulletin, 21(8), 795–805.*

Metallinou, A., Yang, Z., Lee, C., Busso, C., Carnicke, S., & Narayanan, S. (2016). The USC CreativeIT database of multimodal dyadic interactions: From speech and full body motion capture to continuous emotional annotations. *Language Resources and Evaluation, 50(3), 497–521.*

Pallavicini, F., & Pepe, A. (2019). Comparing player experience in video games played in virtual reality or on desktop displays: Immersion, flow, and positive emotions. *Extended Abstracts of the Annual Symposium on Computer-Human Interaction in Play Companion Extended Abstracts, 195–210.*

Pfeiffer, T. (2012). Measuring and visualizing attention in space with 3D attention volumes. *Proceedings of the Symposium on Eye Tracking Research and Applications, 29–36.*

Pfeiffer, T., Latoschik, M. E., & Wachsmuth, I. (2008). Evaluation of binocular eye trackers and algorithms for 3D gaze interaction in virtual reality environments. *JVRB-Journal of Virtual Reality and Broadcasting, 5(16).*

Sáez-López, J. M., Cózar-Gutiérrez, R., González-Calero, J. A., & Gómez Carrasco, C. J. (2020). Augmented reality in higher education: An evaluation program in initial teacher training. *Education Sciences, 10(2), 26.*

Sanders, J. A., & Wiseman, R. L. (1990). The effects of verbal and nonverbal teacher immediacy on perceived cognitive, affective, and behavioral learning in the multicultural classroom. *Communication Education, 39(4), 341–353.*

Smidekova, Z., Janik, M., Minarikova, E., & Holmqvist, K. (2020). *Teachers' gaze over space and time in a real-world classroom. Journal of Eye Movement Research, 13(4).*

Susi, T., Johannesson, M., & Backlund, P. (2007). *Serious games: An overview.*

Volkova, E. P., Mohler, B. J., & Bülthoff, H. H. (2012). Motion Capture of Emotional Body Language in Narrative Scenarios.

Zhonggen, Y. (2019). *A meta-analysis of use of serious games in education over a decade.* International Journal of Computer Games Technology, 2019.

9 Bibliography

References

[1] M. S. El-Nasr and S. Yan, "Visual attention in 3d video games," 2006, pp. 22–es.

[2] D. Steele and R. Zhang, "Enhancement of teacher training: Key to improvement of english education in japan," *Procedia-Social and Behavioral Sciences*, vol. 217, pp. 16–25, 2016.

[3] T. Susi, M. Johannesson, and P. Backlund, "Serious games: An overview," 2007.

[4] K.-F. Kaltenborn and O. Rienhoff, "Virtual reality in medicine," *Methods of information in medicine*, vol. 32, pp. 407–417, 1993.

[5] M. Eskenazi, "An overview of spoken language technology for education," *Speech Communication*, vol. 51, no. 10, pp. 832–844, 2009.

[6] F. L. Cathy Li, "The covid-19 pandemic has changed education forever. this is how," https://www.weforum.org/agenda/2020/04/coronavirus-education-global-covid19-online-digital-learning/.

[7] C. Dede, "Rethinking how to invest in technology," *Educational leadership*, vol. 55, no. 3, pp. 12–16, 1997.

[8] https://www.weforum.org/agenda/2016/05/5-charts-that-explain-the-future-of-education/.

[9] S. Kavanagh, A. Luxton-Reilly, B. Wuensche, and B. Plimmer, "A systematic review of virtual reality in education," *Themes in Science and Technology Education*, vol. 10, pp. 85–119, 2017.

[10] K. Squire, "Video games in education," *Int. J. Intell. Games & Simulation*, vol. 2, no. 1, pp. 49–62, 2003.

[11] J. Hedberg and S. Alexander, "Virtual reality in education: Defining researchable issues," *Educational Media International*, vol. 31, pp. 214–220, 1994.

[12] K. Ahir, K. Govani, R. Gajera, and M. Shah, "Application on virtual reality for enhanced education learning, military training and sports," *Augmented Human Research*, vol. 5, pp. 1–9, 2020.

[13] Y. Zhonggen, "A meta-analysis of use of serious games in education over a decade," *International Journal of Computer Games Technology*, vol. 2019, 2019.

[14] https://dataintelo.com/report/educational-games-market/.

[15] P. Ramsamy, A. Haffegee, R. Jamieson, and V. Alexandrov, "Using haptics to improve immersion in virtual environments," in *International Conference on Computational Science*. Springer, 2006, pp. 603–609.

[16] H. K. Kim, J. Park, Y. Choi, and M. Choe, "Virtual reality sickness questionnaire (vrsq): Motion sickness measurement index in a virtual reality environment," *Applied ergonomics*, vol. 69, pp. 66–73, 2018.

[17] T. Pfeiffer, M. E. Latoschik, and I. Wachsmuth, "Evaluation of binocular eye trackers and algorithms for 3d gaze interaction in virtual reality environments," *JVRB-Journal of Virtual Reality and Broadcasting*, vol. 5, 2008.

[18] Z. Smidekova, M. Janik, E. Minarikova, and K. Holmqvist, "Teachers' gaze over space and time in a real-world classroom," *Journal of Eye Movement Research*, vol. 13, 2020.

[19] E. P. Volkova, B. J. Mohler, and H. H. Bülthoff, "Motion capture of emotional body language in narrative scenarios," 2012.

[20] A. Metallinou, Z. Yang, C. chun Lee, C. Busso, S. Carnicke, and S. Narayanan, "The usc creativeit database of multimodal dyadic interactions: From speech and full body motion capture to continuous emotional annotations," *Language resources and evaluation*, vol. 50, pp. 497–521, 2016.

[21] P. Caserman, A. Garcia-Agundez, R. Konrad, S. Göbel, and R. Steinmetz, "Real-time body tracking in virtual reality using a vive tracker," *Virtual Reality*, vol. 23, no. 2, pp. 155–168, 2019.

[22] C. Girard, J. Ecalle, and A. Magnan, "Serious games as new educational tools: how effective are they? a meta-analysis of recent studies," *Journal of computer assisted learning*, vol. 29, no. 3, pp. 207–219, 2013.

[23] C. Gkantsidis, T. Karagiannis, and M. VojnoviC, "Planet scale software updates," in *Proceedings of the 2006 conference on Applications, technologies, architectures, and protocols for Computer Communications*, 2006, pp. 423–434.

[24] A. Matallaoui, P. Herzig, and R. Zarnekow, "Model-driven serious game development integration of the gamification modeling language gaml with unity," in *2015 48th Hawaii International Conference on System Sciences*. IEEE, 2015, pp. 643–651.

[25] B.-A. Guérin, *ASP. NET con C# en Visual Studio 2017: diseño y desarrollo de aplicaciones Web*. Ediciones ENI., 2018.

[26] A. Wulvik, J. Erichsen, M. Steinert *et al.*, "Capturing body language in engineering design–tools and technologies," *DS 85-1: Proceedings of NordDesign 2016, Volume 1, Trondheim, Norway, 10th-12th August 2016*, pp. 165–174, 2016.

[27] J. A. Edinger and M. L. Patterson, "Nonverbal involvement and social control." *Psychological bulletin*, vol. 93, p. 30, 1983.

[28] J. A. Sanders and R. L. Wiseman, "The effects of verbal and nonverbal teacher immediacy on perceived cognitive, affective, and behavioral learning in the multicultural classroom," *Communication Education*, vol. 39, pp. 341–353, 1990.

[29] F. Khan and H. M. Irshadullah, "A review of the effect of education and good trained teachers on students' performance." *Putaj Humanities & Social Sciences*, vol. 25, no. 2, 2018.

[30] N. A. Stevenson, J. VanLone, and B. R. Barber, "A commentary on the misalignment of teacher education and the need for classroom behavior management skills," *Education and Treatment of Children*, vol. 43, no. 4, pp. 393–404, 2020.

[31] K.-E. Stavroulia and A. Lanitis, "On the potential of using virtual reality for teacher education," in *International Conference on Learning and Collaboration Technologies*. Springer, 2017, pp. 173–186.

[32] L. Darling-Hammond, M. E. Hyler, and M. Gardner, "Effective teacher professional development." 2017.

[33] F. Pallavicini and A. Pepe, "Comparing player experience in video games played in virtual reality or on desktop displays: Immersion, flow, and positive emotions," 2019, pp. 195–210.

[34] T. Pfeiffer, “Measuring and visualizing attention in space with 3d attention volumes,” 2012, pp. 29–36.

[35] Unity2022.2, “Unity unity documentation, processing,” https://docs.unity3d.com/2022.2/Documentation/Manual/dynamic-batching.html.

[36] Unity, “Unity unity documentation, lightning,” https://docs.unity3d.com/Manual/LightModes.html.

[37] S. Kiesler and J. N. Cummings, “What do we know about proximity and distance in work groups? a legacy of research,” *Distributed work*, vol. 1, pp. 57–80, 2002.