Université de Montréal

Theseus: A 3D Virtual Reality Orientation Game with Real-time Guidance System for Cognitive Training

par

Manish Kumar Jha

Département d'informatique et de recherche opérationnelle
Faculté des arts et des sciences

Mémoire présenté en vue de l'obtention du grade de
Maître ès sciences (M.Sc.)
en Informatique

October 21, 2020

Université de Montréal

Faculté des arts et des sciences

Ce mémoire intitulé

Theseus: A 3D Virtual Reality Orientation Game with Real-time Guidance System for Cognitive Training

présenté par

Manish Kumar Jha

a été évalué par un jury composé des personnes suivantes :

Esma Aïmeur
(président-rapporteur)

Claude Frasson
(directeur de recherche)

Sébastien Roy
(membre du jury)

Résumé

Des études soutiennent que l'entraînement cognitif est une méthode efficace pour ralentir le déclin cognitif chez les personnes âgées. Les jeux sérieux basés sur la réalité virtuelle (RV) ont trouvé une application dans ce domaine en raison du haut niveau d'immersion et d'interactivité offert par les environnements virtuels (EV). Ce projet implémente un jeu d'orientation 3D en réalité virtuelle entièrement immersif avec un système pour guider l'utilisateur en temps réel. Le jeu d'orientation 3D est utilisé comme exercice pour entraîner les capacités cognitives des utilisateurs. Les effets immédiats du jeu d'orientation sur les capacités de mémoire et d'attention ont été étudiés sur quinze personnes âgées présentant un déclin cognitif subjectif (DCS). Il a été observé que bien qu'il n'y ait pas eu d'amélioration significative des résultats pour les exercices d'attention, les participants ont obtenu de meilleurs résultats aux exercices de mémoire spécifiques après avoir joué au jeu d'orientation.

Le manque de succès dans la réalisation de l'objectif requis peut parfois augmenter les émotions négatives chez les êtres humains, et plus particulièrement chez les personnes qui souffrent de déclin cognitif. C'est pourquoi le jeu a été équipé d'un système de guidage avec indices de localisation en temps réel pour contrôler les émotions négatives et aider les participants à accomplir leurs tâches. Le système de guidage est basé sur des règles logiques; chaque indice est délivré si une condition spécifique est remplie. Le changement des émotions des participants a montré que les indices sont efficaces pour réduire la frustration, étant donné qu'ils sont facilement compréhensibles et conçus pour donner un retour positif.

La dernière partie du projet se concentre sur le système de guidage et met en œuvre un moyen pour l'activer entièrement selon les émotions d'une personne. Le problème consiste à identifier l'état des émotions qui devraient déclencher l'activation du système de guidage. Ce problème prend la forme d'un processus de décision markovien (PDM), qui peut être résolu via l'apprentissage par renforcement (AR). Le réseau profond Q (RPQ) avec relecture d'expérience (ER), qui est l'un des algorithmes d'apprentissage par renforcement les plus avancés pour la prédiction d'actions dans un espace d'action discret, a été utilisé dans ce contexte. L'algorithme a été formé sur des données d'émotions simulées, et testé sur les données de quinze personnes âgées acquises lors d'expériences menées dans la première partie du projet. On observe que la méthode basée sur l'AR est plus performante que la méthode basée sur les règles pour identifier l'état mental d'une personne afin de lui fournir des indices.

Mots-clés : Maladie d'Alzheimer, Émotions, Réalité virtuelle immersive, Orientation spatiale, Indices, Apprentissage par renforcement, Réseaux profond Q.

Abstract

Studies support cognitive training as an efficient method to slow the cognitive decline in older adults. Virtual reality (VR) based serious games have found application in this field due to the high level of immersion and interactivity offered by virtual environments (VE). This project implements a fully immersive 3D virtual reality orientation game with a real-time guidance system to be used as an exercise for cognitive training. The immediate aftereffects of playing the orientation game on memory and attention abilities were studied on fifteen older adults with subjective cognitive decline (SCD). It was observed that while there was no significant improvement in attention exercises, the participants performed better in specific memory exercises after playing the orientation game.

Sometimes lack of success in achieving the required objective may increase the negative emotions in humans and more so in people who suffer from cognitive decline. Hence, the game was equipped with a real-time guidance system with location hints to control negative emotions and help participants to complete the tasks. The guidance system is based on logical rules; each hint is delivered if a specific condition is met. Change in emotions of participants showed that hints are effective in reducing frustration, given that the hints are easily comprehensible and designed to give positive feedback.

The final part of the project focuses on the guidance system and implements a way to activate it entirely based on a person's emotions. The problem calls for identifying the state of the emotions that should trigger the guidance system's activation. This problem takes the form of a Markov decision process (MDP), which can be solved by setting it in a reinforcement learning framework. Deep Q-Learning network (DQN) with experience replay (ER), which is one of the state-of-the-art reinforcement learning algorithms for predicting actions in discrete action space, was used in this context. The algorithm was trained on simulated data of emotions and tested on the data of fifteen older adults acquired in experiments conducted in the first part of the project. It is observed that the RL based method performs better than the rule-based method in identifying the mental state of a person to provide hints.

Keywords: Alzheimer's Disease, Emotions, Immersive Virtual Reality, Spatial Orientation, Hints, Reinforcement Learning, Deep Q-Learning Networks.

Contents

List of Tables

List of Figures

List of acronyms and abbreviations

ACER	Actor-Critic with Experience Replay
AD	Alzheimer's Disease
CT	Cognitive Training
DQN	Deep Q-learning Network
ECG	Electrocardiogram
EEG	Electroencephalogram
ER	Experience Replay
HMD	Head-mounted display
MCI	Mild Cognitive Impairment
MDP	Markov Decision Process
MSE	Mean Square Error

RL	Reinforcement Learning
SCD	Subjective Cognitive Decline
VE	Virtual Environment
VR	Virtual Reality

Acknowledgment

I express my sincere gratitude to my supervisor, Professor Claude Frasson, who guided and encouraged me through the course of this project. Without his persistent help, the goal of this project would not have been realized.

I acknowledge the National Science and Engineering Research Council, Canada (NSERC), and BMU Labs for believing in this project and funding this work.

I thank my colleagues Hamdi Ben Abdessalam and Marwa Boukadida who collaborated with me on this project and helped me with useful discussions and suggestions and provided their useful insights whenever required.

Chapter 1

Introduction

1.1. General Context

Alzheimer's disease (AD) is an irreversible, progressive form of dementia[1] that negatively affects memory, thinking, and behaviour. The early symptoms of AD are memory loss, lack of attention, and spatial disorientation, to name a few. These problems substantially affect the quality of life of a person and can lead to other symptoms such as challenges in planning tasks, poor judgment, withdrawal from social activities and surge in negative emotions like depression and anxiety. It mostly affects older people over the age of 65 and accounts for 60-70 percent of all the cases of dementia [**2**]. Since there is no known cure for Alzheimer's Disease(AD), current approaches emphasize on helping people preserve mental function and control behavioural symptoms. A significant number of scientific researchers have focused on non-pharmacological interventions to slow down the progression of the disease. Cognitive training is one such approach.

Cognitive training(CT) or brain training refers to a set of theoretically motivated mental exercises often in the form of games, puzzles, or tasks related to daily activities that require a person to employ particular mental ability to achieve the objectives in order to optimize the cognitive functioning. Several studies show that the human brain remains plastic throughout life, continuously evolving and adapting as we learn new things and experience different situations in life [**3, 4, 5**]. This property of brain known as neuroplasticity lends support to the basic principle of cognitive training that mental abilities can be upheld and improved by exercising the brain, similar to how physical fitness can be enhanced by exercising the body. Exercises in CT sessions target abilities such as working memory, attention, reasoning, rule acquisition, spatial orientation, and other higher cognitive functions. These exercises are designed so that the learning effects of training sessions

[1]Dementia is the general term for a decline in mental ability severe enough to interfere with daily life.

can be transferred to the cognitive functions in daily life. The underlying hypothesis is that transfer of learning can occur if the training exercises and transfer tasks engage specific overlapping neural components [**6**]. This type of learning is known as near-transfer, where the same brain regions are active in a training exercise and daily life tasks. A study on cognitive training in older adults by Ball et al. [**7**] showed that training specific skills in memory and attention problem solving enhanced the particular skills in which they were trained . In the case of older adults experiencing cognitive decline, the goal extends to slow down the overall cognitive decline and not only particular mental abilities [**8**]. Some research work suggest that transfer to general executive control functions is possible if the training model is flexible, employing an array of variable tasks and conducted regularly over time [**8, 9, 10, 11**]. Effects of cognitive training has been studied not only in the case of AD, but also in association with other cognitive impairment disorders as well. Kristen et al. showed that cognitive training can increase the ability of patients to activate the pre-frontal cortex regions subserving attention and working memory [**12**]. Thus, cognitive training can be used to improve the cognitive skills at which the training program is targeted and generalize the improvement to functional abilities useful in daily life.

As discussed earlier, training programs can be designed to improve different executive functions, in a general manner or to target a specific skill. Spatial orientation is one of the cognitive abilities that decline with ageing [**13**]. Disorientation is one of the early symptoms of Alzheimer's Disease as people with AD exhibit a significantly higher decline compared to normal ageing. Lovden et al.[**14**] in their research showed that spatial navigation training programs in older men led to improvements in spatial performances, suggesting that spatial navigation experience may protect against age-related structural changes in the brain and associated cognitive decline. It is believed that the formation of a cognitive map is critical to orientation skills [**15**]. According to the cognitive map theory, the formation of representations of spatial information using spatial reference frames and cues from the surroundings (also known as the creation of a cognitive map), helps reduce cognitive load and increases recall and the encoding of novel information [**15, 16**]. It has been observed that a decrease in the volume of the hippocampus[2], correlates with a decline in cognitive function. It is hypothesized that increasing grey matter in the hippocampus could entail better memory [**17**]. A recent study by West et al. [**18**] showed that playing 3D video games over a period, such as Super Mario 64, reportedly increases hippocampal volume. Our research is based on these findings to examine if a 3D virtual reality based orientation game can improve memory and attention capabilities in addition to spatial orientation skills.

[2]Hippocampus is the part of brain which plays an important role in the consolidation of information from short-term memory to long-term memory, and in spatial memory that enables navigation.

1.1.1. Virtual Reality in Cognitive Training

Virtual reality can be defined as 'real-time interactive graphics with 3D models, combined with a display technology that gives the user the immersion in the model world and direct manipulation' [**19**]. The earliest VR application dates back to 1962, when Morton Heilig, one of the pioneers of virtual reality(VR) applications, patented 'a telescopic television apparatus for individual use'. The primary objective behind the development of the Sensorama simulator was 'to teach and train individuals without actually subjecting the individuals to possible hazards of particular situations' [**20**]. The simulator looked like an arcade game set and gave the user an experience of virtually riding a bike through the streets of Brooklyn. In addition to a 3D view of the city, the user could feel the breeze, the odour and the vibrations or jolts similar to the movement of a bike in the real life. The product was not well-received in the business community owing to high production costs. Fast forward more than half a century, today VR technologies are being used in a multitude of applications such as gaming, military training, pilot training, cinematography, architecture design, healthcare, science, and education.

Today's advanced VR systems usually consist of a display equipped with audio-visual outputs, multiple sensors to track head(sometimes eyes) and optionally body movements, and accept feedback via user input devices such as joystick. The interactive simulated model of the world inside a VR application is traditionally called a virtual environment(VE). Peitro et al. [**21**] discuss three characteristic features of VR systems in their review work of research on virtual and augmented reality: immersion, sense of presence and interaction with the virtual environment. Immersion refers to what a VR system imparts from a technical point of view (field of view, frame rate, latency, stereo audio) in relation to the degree of 'fidelity' to the real-world equivalent of the virtual environment [**22**]. Virtual environments can be categorized broadly as non-immersive, semi-immersive and fully-immersive [**23**]. Both non-immersive and semi-immersive VR applications use desktops to display the environment, and the primary difference between the two being in the level of detail and graphics quality. A fully-immersive VR system consists of a head-mounted display(HMD) and offers more fidelity to the dimensions and details of the objects and 3D models within the virtual environment. The sense of presence is the human reaction to immersion [**22**]. Contrary to immersion, presence is a subjective concept and may vary from person to person in the same virtual environment. However, research shows that presence and immersion are correlated as an increase in the degree of immersion induces a higher sense of presence [**24**].

While VR certainly brings the reality closer in a simulated environment, in some cases it leads some users to develop an affliction similar to motion sickness. In VR environments with components involving user displacement with respect to seemingly stationary reference frames(such as trees, land) induces an ailment called cybersickness in some individuals which has similar symptoms (headache, dryness of mouth, nausea, vertigo) as motion sickness. Though the exact causes are attributed a multitude of factors, it is postulated that while motion sickness is caused by vestibular[3] and visual stimulation, cybersickness can be caused by the visual stimulation alone [**25**]. This limits the usability of the VR environment for training purposes to some degree, however, several approaches are suggested to reduce the stimulation leading to this problem. The use of motion platform [**26**] aligned with visual input could help reduce cybersickness. Another suggested method is direct vestibular stimulation by using a device that sends electrical signals to the brain and tricks the vestibular system into believing that motion is taking place in the real-world [**25**]. Weech et al. [**27**] noted that several studies support that presence and cybersickness in VR are negatively correlated which means that the more fidelity of the environment to reality, the lesser will be the risk of the sickness.

Virtual reality-based cognitive training offers many benefits given that the associated problem of cybersickness is addressed. A high degree of immersion in VR environments promote high sense of presence and quality interaction, thus playing a crucial role in stimulating the senses and holding the person's attention for a more extended period in training sessions. Game engines such as Unreal and Unity 3D allow the creation of highly engaging and diverse virtual environments that can offer training tasks ranging from indoor tasks such as cooking and shopping to outdoor tasks such as driving a vehicle or riding a train [**7, 28, 29, 30**]. These training sessions can be conducted safely indoors anywhere where the patients feel comfortable, thus avoiding any risk posed otherwise by a real-life environment. Virtual reality-based training sessions can be supplemented with equipments to facilitate the collection of physiological data such as electroenceophalography(EEG)[4] and electrocardiography(ECG)[5], which can be converted into input control to adapt the elements of the virtual environment and the training tasks in real-time [**31**].

Over the years, many studies have implemented multiple VR environments for cognitive training. In one of the earlier studies, Rizzo et al. [**32**] observed attention enhancement in young students in a VR classroom compared to the non-VR control group. Optale et al. [**33**] observed that VR memory training during a period of three months presented an improvement in long term memory in elders with memory impairment. Gamito et al. [**34**] used working memory tasks (i.e. buying several items), visuospatial orientation tasks (i.e.

[3]The vestibular sensory system keeps the brain informed about motion, head position, and spatial orientation; and allows us to stabilize our head and body during movement and maintain our balance and posture.
[4]EEG is used to record electrical activity of the brain by use of non-invasive electrodes placed on head.
[5]ECG is used to record electrical activity of the heart and provides information about heart rate and rhythm.

finding the way to the minimarket) in a VR environment for the training of stroke patients two to three sessions per week over the four to six week period of treatment. Their study revealed a general improvement in attention and memory abilities over multiple training sessions. Thus VR training programs have proved to be promising in enhancing cognitive abilities.

1.1.2. Help System in Games

For the elderly, continuous feedback and guidance to complete the quests is one of the main ingredients that must be considered while designing the cognitive training tasks. Games used for training purpose that have elements which encourage positive emotions and tense feelings, rather than excitement and continuous challenge are more suited to the elders [**35**]. As per a recent study by Caggianese et al. [**36**], it is more beneficial for AD patients to be helped through the completion of a challenge rather than having them to give up. It is vital to present both audio and visual cues to cater to the needs of a specific profile of patients suffering from either visual or auditory impairments [**33**]. Thus, real-time assistance with audio and visual feedback is one of the mandatory components in games for elders, to incorporate mechanism which achieves high-level engagement by keeping player filled with positive emotions. Fisch in his work [**37**] on educational computer games has suggested that the hint in games should complement the gameplay and the learning, leading the person in the right direction to discover the answer themselves, instead of revealing the correct answer. Furthermore, he suggested that the help provided should be leveled - meaning the second hint provided should be more detailed than the first one and so on. The underlying argument is that requesting the next hint indicates the fact that the student is still stuck after receiving the current hint and need more details to solve the problem.

1.1.3. Emotion Based Feedback in Learning

Emotions are the purest form of expressions as one can express their emotions more conveniently than words or actions. Over the years, several approaches to measure the emotions have been developed such as physiological sensing approaches like electroencephalography (EEG), face tracking [**38**] and eye tracking [**39**] which are coupled with computer based assessments. Kiel et al. [**40**] defined frustration in games as the negative emotion which arises when the progress of a user towards achieving a goal in the game is impeded. Frustration signifies that the player is in need of assistance to achieve the objective. Engagement signify how much interest the player feels in interacting with the elements of the game. EEG to detect and analyze mental and emotional states to assist in learning by playing intelligent video games by placing a headgear while the player is interacting with the game. Ghali et al.

[**41**] used EEG signals to assess participants' mental states and focused on their engagement and frustration and proposed help strategies in a physics game.

1.2. Current Work

People with early-stage stage Alzheimer's Disease experience particular difficulty in processing new information and forming memories, but rates of forgetting are not yet elevated [**42**]. Interventions that can help with memory difficulties may have the potential to further reduce secondary problems such as emotional imbalance and improve the overall well-being of the person. A remote goal in the field of cognitive training is a training program which can be easily administered regularly by a person or their caregiver themselves at home, without the need of travelling for such an exercise. To fulfill this goal, several fronts have to be covered. First, we would require such training programs which would not only improve the cognitive abilities which they train, but also transfer the benefits to other abilities. Second, we would need accurate devices to measure EEG and/or emotions accurately. And finally, we need feedback systems which make the training adaptable the profile of individual person and cater to their needs rather than having one generalised training program. This work is a small step covering the first and third front of this goal.

1.2.1. Research Objective

In this thesis, we present a fully immersive 3D virtual environment (VE) called Theseus, where the participant must find listed items in order in a garden using his/her orientation skills. We implemented a rule-based guidance system that helps the participant look for the items in the environment. It assesses the emotions of the participants and keeps track of their location in real-time in order to provide hints without being asked explicitly for any help. Our goal for this study was to investigate the effect of the orientation game and the guidance system on the cognitive functions of people suffering from subjective cognitive decline (SCD), a pre-clinical state of possible Alzheimer's Disease (AD). We hypothesize that VR based orientation games which can induce the generation of cognitive maps could positively enhance attention and memory abilities in people with cognitive decline.
We state our research questions as the following:

(1) Is it possible to stimulate the brain using a virtual maze game in order to enhance attention and memory?
(2) Does being helped by the guidance system in the cognitive tasks effectively reduce frustration?

1.3. Thesis Organization

This thesis discusses the implementation details of Theseus, the orientation game and guidance system, the experiment methodology and results. Chapter 2 explores the related work in the field of cognitive training using virtual reality and transfer of the learning abilities to tasks other than the training tasks.

Chapter 3 provides the implementation details of the Orientation Game and the objectives of the game. It explains the design of the game and the featured elements. Next, it discusses the details of the guidance system: underlying assumptions, different level of hints, criteria for activation and change in the level of hints.

Chapter 4 describes the methodology of the experiments conducted. It discusses the Emotive EPOC EEG device and Fove VR headset used in our experiments. Then, we explain the different attention and memory exercises which were conducted before and after the orientation game to assess change in performance of the participants. Finally, the complete process of the experiments is illustrated.

Chapter 5 highlights the important observations. It discusses the performance changes in memory and attention exercises of each participant after playing the orientation game. We also evaluate the usage of the hints provided by the guidance system and it's effect on the emotions of the participants.

Chapter 6 summarises this dissertation and discusses the directions for future work.

Appendix A states the paper 'Improving Cognitive and Emotional State Using 3D Virtual Reality Orientation Game' presented at 16th International Conference on Intelligent Tutoring Systems. Appendix B displays the VR scenes of the objectives of the game that is to collect four items of interest in order. Appendix C gives the comparison of hints provided by the rule-based system with reinforcement learning based guidance system. The questionnaire after the experiments to be completed by the participants is listed in Appendix D.

Chapter 2

Theseus: The Orientation Game

2.1. The Environment

The orientation game is based on the motivation that creation of cognitive maps using spatial reference in the brain during navigation helps to improve the functions of the hippocampus region. The environment of the game simulates a botanical garden in the form of a 5x5 maze. In this environment, trees constitute the maze walls and clearings through the trees form the pathways. The game starts at one corner of the garden, and a person can easily navigate using a joystick by clicking in the intended direction to move forward. Since the experiments were conducted with the local population of Montreal (QC), the audio and visual messages and other text elements of the game are in the French language. Language can be customized easily, depending on the target audience. Different elements in the game are:

- Map with geographical directions
- Position and direction of the user show by a red arrow in map
- List of items to be collected in order
- Location of next item shows as blue circle
- Visual messages (tasks related commands or hints)
- Audio messages (tasks related commands or hints)

Before the actual game starts, a tutorial is provided to familiarize the users with controls, navigation, tasks and different elements of the game. The participant is given an example of an item to look for and collect, which allows them to understand the movement within the environment. Next, the different types of hints provided by the guidance system are explained. The game starts once the user is familiar with the controls, navigation, and tasks. Figure 3.1 shows a screenshot of the start screen of the game where games' objectives are conveyed to the participant. The participant needs to collect three items available at specific locations in the virtual environment in a given order and put them back in a basket available at the start of the location. The user is shown the list of items and the

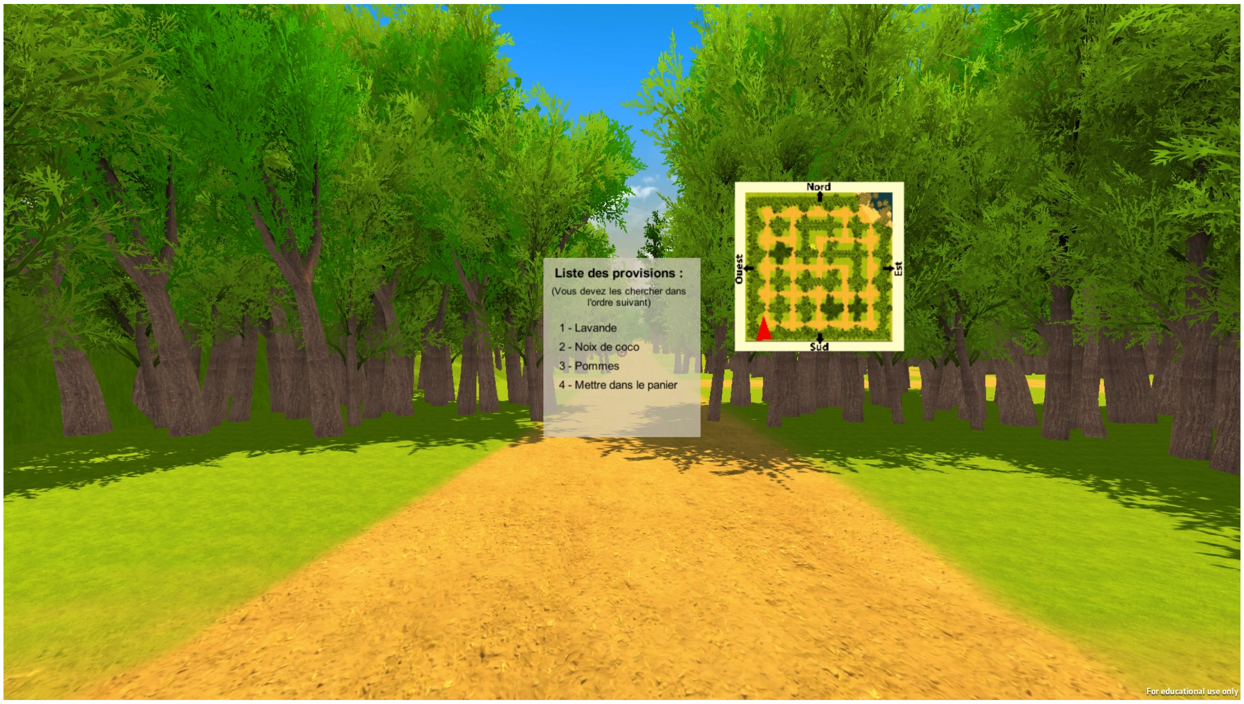


Figure 2.1. Environment showing list of items, map, and location of user

location of the first item in the environment by a flashing blue circle on the map. The blue circle is displayed for five seconds, after which it vanishes from the map. The user needs to memorize the location of the item to collect it. As soon as the item is collected, we delete it from the list and show the next item. This sequence continues until the user collects all the items. For the last task, the user is asked to return to the starting point and put the three collected items in a basket. To minimize the risk of cybersickness, the animation of motion from one point to another was skipped and the users were simply moved to the new location at the click of the joystick. The change of reference frames and the movement of red arrow to display the position of user in the map conveyed the movement within the virtual environment.

2.2. The Guidance System

The guidance system is a rule-based system that provides navigational hints or audio and visual messages to the participants and helps them in completing the quests. It actively monitors the participant's emotions using Emotiv electroencephalograph (EEG) headset data in real-time. It also keeps track of the participant's movement and actions that are taken while completing the tasks. The guidance system sits outside the VR environment and the emotions tracking system and receives the person's emotion and movement data every second. Frustration was selected as the negative emotion as it can signify if the player

Hint levels		Human's position	Object's location	Message (Audio and Visual)	Highlighted Path in the Map
Level 1		✓	✓		
Level 2	2-1	✓	✓	Please check the position of object on the map and try to reach it.	
	2-2	✓	✓	You are too far. Try to take few steps back.	
	2-3	✓	✓	1. Good job! Almost there. 2. Keep up the good work 3. You are going in the right direction	
Level 3		✓		Follow the displayed path. The object is somewhere nearby.	Path to cell nearest to target object
Level 4		✓		Follow the given path to find the object.	Complete path to the object.

Table 2.1. Different level of hints.

is in use of assistance[**40**] and engagement and excitement as the positive emotions. The emotion values are furnished by Emotiv emotion detection suite. Apart from frustration, confusion is another emotion which people sense when disoriented. Frustration was chosen over confusion direct measure of confusion via Emotiv Epoc is not available and confusion and inability to solve task lead to frustration. It evaluates the change in emotions and movement of the user every fifteen seconds to assess if the user is experiencing high frustration or is lost. The user is considered lost or *'away from target'* in the environment if he/she is not able to locate the required item after certain number of steps irrespective of their emotion levels. This may be caused due to unfamiliarity with such VR systems and/or because of difficulty in understanding the elements of the orientation game. On sensing a situation where the participant may need a hint, it sends a message to the VR system, which displays the hints in the form of location in the map or audio and visual messages in case of text-based hints.

The guidance system has four levels of hints, each of which provide progressively more information to find the items compared to the previous level. Table 2.1 displays the level of detail provided in different hints. Level 1 provides the least information and displays only the participant's location and the object's location on the map as a flashing blue circle. In subsequent level of hints, the hint in map is accompanied by an audio-visual message. There are three different scenarios in second level of hints. If the user is wandering too far

Figure 2.2. Different level of hints as displayed in the environment. In level 1 and level 2 hints, blue circle represents the target location. In level 3 hint and level 4 hint, the target location is at the upper left corner of the map.

from the object to be collected, level 2-1 asks the user to verify the location of object in the map. Further if the user experiences high frustration and is going away from the object, level 2-2 is activated which warns the user to take few steps back as they might have been closer to the object earlier. If the user experiences rapid increase in frustration in spite of going on correct path, level 2-3 gives an encouraging message. If the user cannot find the object after level 2 hints, more information is provided in next levels of hints. Level 3 hint highlights a path in the map, which the participant can follow to reach a location that is close to the actual location of the object. This still leaves some scope of exploration and allows the participant to look for the object in all possible directions. Level 4 hint highlights the complete path leading to the object's location on the map. Both level 3 and level 4 hints are accompanied by audio-visual messages. Figure 2.2 shows the different hints as they appear in the virtual reality environment.

The hints provided by the guidance system are activated in three different cases:

(1) **Emotions**: At every timestamp, the mean of the change and the rate of the change of emotions are used to calculate a net score for each emotion. The emotion with the

maximum score is compared with an empirically defined threshold. A score higher than the threshold activates the emotion-based hints.

(2) **Away from target**: If the participant takes three steps or more, all of which are at four blocks or more from the target, the hint level 2-1 is activated.

(3) **No Movement**: If the participant does not move for more than a given amount of time, a hint based on no movement is activated.

As discussed earlier, the amount of details provided by the guidance system increases progressively with the level of hints. At first, the criteria to activate hints by the case(2) and (3) mentioned is examined and if they are satisfied, the appropriate level of hint is presented. When the participant is considered away from the object as determined by the guidance system, the hint level 2-1 is rendered. In case of hint activated due to no movement of the participant, the level increases every fifteen seconds to the highest of level 4 hint. When the participant eventually moves, the hint level resets to 1.

In case of activation based on participant's emotion, if the hint is triggered by frustration, the level of hint increases which subsequently provides more details to find the object. On the other hand, if the hint is triggered by positive emotions such as excitement or engagement, the level of hint decreases. Algorithm 2.2.1 summarises the activation and change in level of hint using emotions. For each of the emotions: frustration, engagement and excitement, the mean of the emotions and rate of change in over past $2t_1$ seconds($t_1 = 15$ in this project) is calculated. Based on these values, an $emoti\_score$ was calculated. If the emotion with highest $emoti\_score$ was frustration, the level of hint would increase every t_1 seconds to a maximum of four. In other cases(excitement or engagement), the hint would decrease to a minimum of zero. Time without hint was reset to -1 every time a hint was provided, else it would increase every second.

Algorithm 2.2.1. Activating and changing level of hints using emotions.

Initialize t_l = hint time interval, $df_{emotion}$ = dataframe to store past emotions,
Initialize time without hint = 0, hint level = 0
While not gameOver:
    Receive emotion values and store in $df_{emotion}$
    If human is not at target location:
        If time without hint > t_l:
            For each emotion:
                $\mu_{emotion}(t : t - t_l)$ = mean of emotions in last t_l seconds
                $\mu_{\Delta emotion}(t : t_l)$ = mean rate of change of each emotion in last t_l seconds
                $\mu_{emotion}(t - t_l : t - 2t_l)$ = mean of emotions in the range of last $t_l - 2t_l$ seconds
                $\mu_{\Delta emotion}(t - t_l : t - 2t_l)$= mean rate of change of emotions in the range of last $t_l - 2t_l$ seconds
                $emoti\_score = 0.5 * (\mu_{\Delta emotion}(t : t_l) - \mu_{emotion}(t : t - t_l)$
                $+\mu_{\Delta emotion}(t - t_l : t - 2t_l) - \mu_{emotion}(t - t_l : t - 2t_l))$
        if emotion with max($emoti\_score$) is frustration
            Increase hint level by 1 to a maximum of 4
        else
            Decrease hint level by 1 to a minimum of 0
        time without hint = -1
    time without hint += 1

In this chapter, we discussed the Orientation game and the guidace system. In next chapter, we will discuss the process of the experiments.

Chapter 3

Method of the Experiment

3.1. Apparatus

3.1.1. Emotiv Epoc

Emotiv EPOC wireless EEG headset (figure 3.1(a)) is an inexpensive commercially available device designed for advanced brain-computer interface applications. It is a fourteen channel headset based on internationally recognized 10-20 EEG system [**43**] (as shown in figure 3.2) and returns the value of five principal frequency bands of brain waves: delta (0.5–4 Hz), theta (3–7 Hz), alpha (8–12 Hz), beta (13–30 Hz) and gamma (30–40 Hz). Rytis et al. [**44**] showed in their research that among several consumer grade EEG devices, Emotiv EPOC performed better and may be more suitable for control tasks using attention/meditation level. The device is furnished with an emotion detection suite that gives emotion values of frustration, engagement, excitement, meditation and valence normalized in the range 0-1. Several studies have established the reliability of these outputs [**45, 46, 47**].

3.1.2. Fove VR

Fove VR (figure 3.1(b)) is an advanced lightweight virtual reality headset equipped with infrared eye-tracking, orientation tracking and position tracking systems. It can be used with VR environments developed with Unity 3D using the Fove Unity plugin. It comes with a WQHD OLED display with a frame rate of 70 fps and provides 100° field of view providing a high level of immersion in the virtual environment.

3.2. Attention and Memory Exercises

Our main objective is to analyze the benefits of the orientation game and cognitive map formation on the attention and memory performances. As discussed earlier, it has been postulated that spatial navigation training programs in older men led may protect against

(a) Emotive EPOC headset (b) Fove VR headset

Figure 3.1. Equipments used to collect EEG and display VR environment

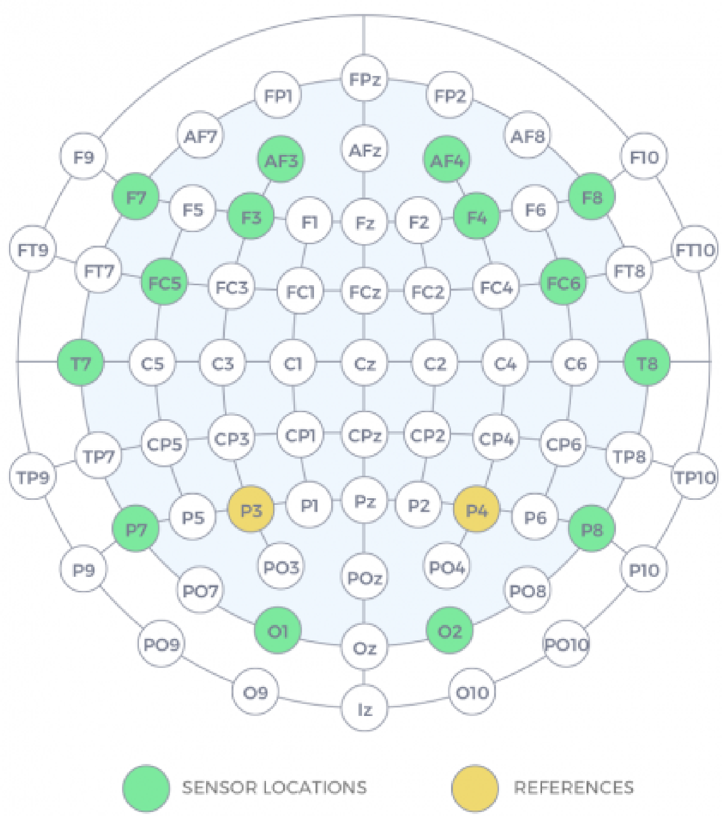


Figure 3.2. 10-20 EEG System with Emotiv EPOC EEG electrode placement highlighted

age-related structural changes in the brain and associated cognitive decline. The creation of cognitive maps in orientation tasks using spatial references and cues require coding of new memory in the brain which may help to improve performance in memory. Also, to achieve the objectives of the game and complete all the tasks the player has to concentrate their attention on the game elements and use them to accomplish the tasks. Therefore, we designed three attention and three memory exercises to compare participants' performance before and after the training game. The attention and memory exercises developed for the experiments are inspired by Montreal Cognitive Assessment (MOCA), a cognitive screening test designed to assist health professionals in detecting mild cognitive impairment and Alzheimer's disease [**48**]. All the exercises are computer-based unlike the orientation game itself which is available in 3D VR system using a head mounted display. This is to assess if the orientation based 3D games could have any impact on general memory and attention outside the virtual reality.

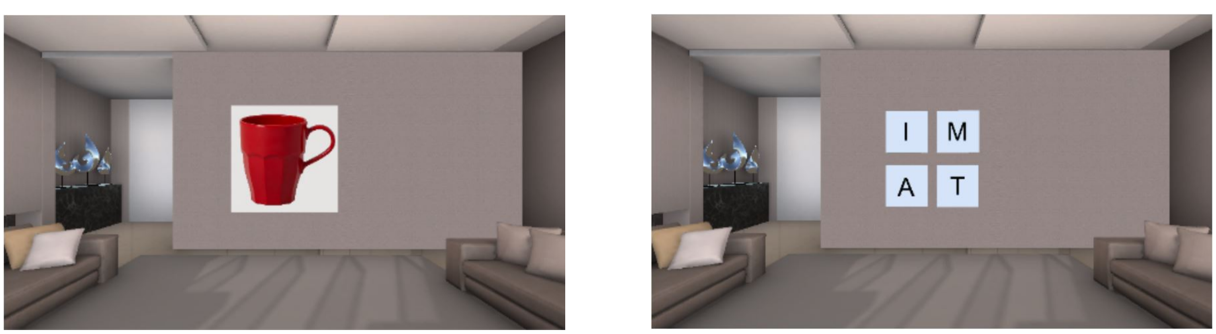
3.2.1. Attention Exercises

The three attention exercises are as follows :

(1) First, a sound sequence of five numbers is played to the participants. The participants have to repeat the same sequence on a numerical pad displayed afterwards. Again, another sound sequence of 3 numbers is played, and this time participants need to report the sequence in the reverse order.

(2) A sound sequence of different letters is played at a rate of one letter per second. The participants have to hit the space bar every time they hear the letter "A."

(3) In the third attention exercise, an object's image is displayed for four seconds on the computer screen. After four seconds, the image is replaced by four letters. The participants need to select the correct letter, which is the initial letter of the name of the object displayed earlier. Figure 3.3 shows an example of this exercise. An image of a mug ("Tasse" in French) is presented. Among the answer choices, the participants must choose the letter "T" for "Tasse".



(a) Displayed Object (b) Answer Choices

Figure 3.3. Third attention exercise

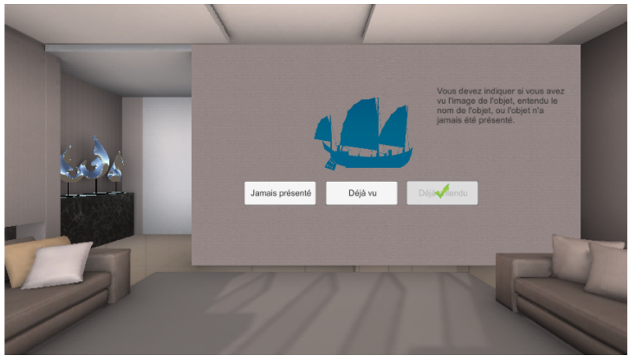


Figure 3.4. Exercise for contextual memory

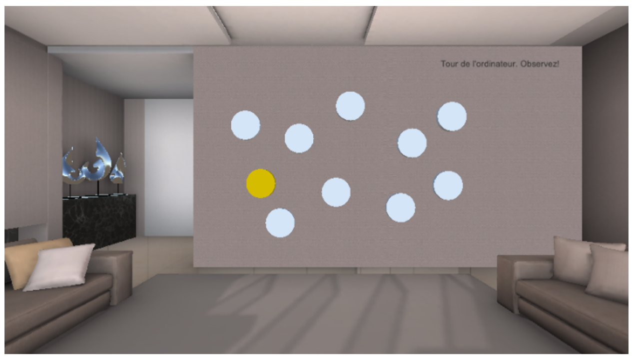
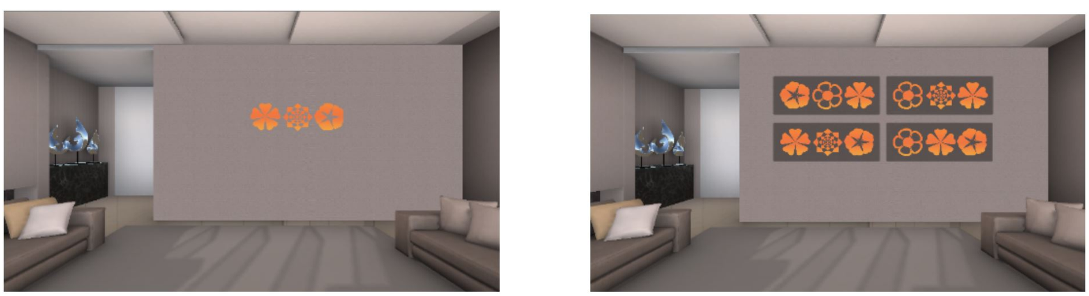


Figure 3.5. Short-term memory exercise



(a) Displayed Object

(b) Answer Choices

Figure 3.6. Exercise for working memory

3.2.2. Memory Exercises

The three memory exercises are as follows:

(1) First memory exercise is a test of contextual memory. The participant is asked to memorize a sequence of objects that are either visually presented on the screen or orally played through the speakers. In the next step, another sequence of different objects is shown on screen, or their names are played. For each object presented in the second sequence, participants have to determine whether the object was visually

shown or played or not presented at all. Figure 3.4 shows the example of an object shown in the second sequence, where the participants select 'Déjà entendu'('already heard') among the three options mentioned earlier.

(2) Second memory exercise is a short-term memory test. On the screen, ten white circles are shown at first. Among the displayed circles, some of them are highlighted one by one in a sequence. The participant is then asked to remember and reproduce the same sequence by clicking the correct circles in order. This exercise is performed twice with different levels of difficulty. Figure 3.5 shows an example of short-term memory exercise.

(3) The last exercise of this series is aimed at testing working memory. A set of 3 pictures is displayed for a short period, and the participant is asked to memorize the set. In the next screen, the participant is asked to identify the set presented among the four displayed answer choices. This exercise is conducted twice with two different sets of pictures. Fugure 3.6 gives an example of a picture set and the answer choices.

3.3. Process of the Experiments

All the experiments were conducted in collaboration with Institut universitaire de gériatrie de Montréal (IUGM) who invited fifteen participants (mean age = 73.4, SD =5.73) with subjective cognitive decline (SCD) to take part in the training sessions. The entire experiment was divided into two sessions. In the first session, the format of the study and the details of the experiment process were explained to the participants. They were invited to sign a consent form and perform some clinical tests to confirm diagnosis and determine whether they were eligible to participate in the experiment. If the participants were deemed eligible to participate in experiments, they were invited to fill three pre-experiment forms: the Positive and Negative Affect Schedule (PANAS) scale [**49**], a self-assessment of emotions, and a questionnaire on cyber-sickness [**50**].

After completing the pre-assessment forms, the participants took part in the experimental session. They were furnished with an Emotic Epoc headset and asked to solve the attention and memory exercises. Once they finish all the exercises, they were allowed to relax for a few minutes before being equipped with the Fove VR headset. With VR headset on, they went through the tutorial of the orientation game. Once they were familiar with the game's control and objectives, they could start playing the orientation game. There was no time limit for completing the tasks in the game. Each participant plays the orientation game once. After completing all the tasks, the Fove VR headset was removed, and participants could relax for a couple of minutes. Once they were ready, they were asked to carry out similar attention and memory exercises with different examples. After completing the exercise, the

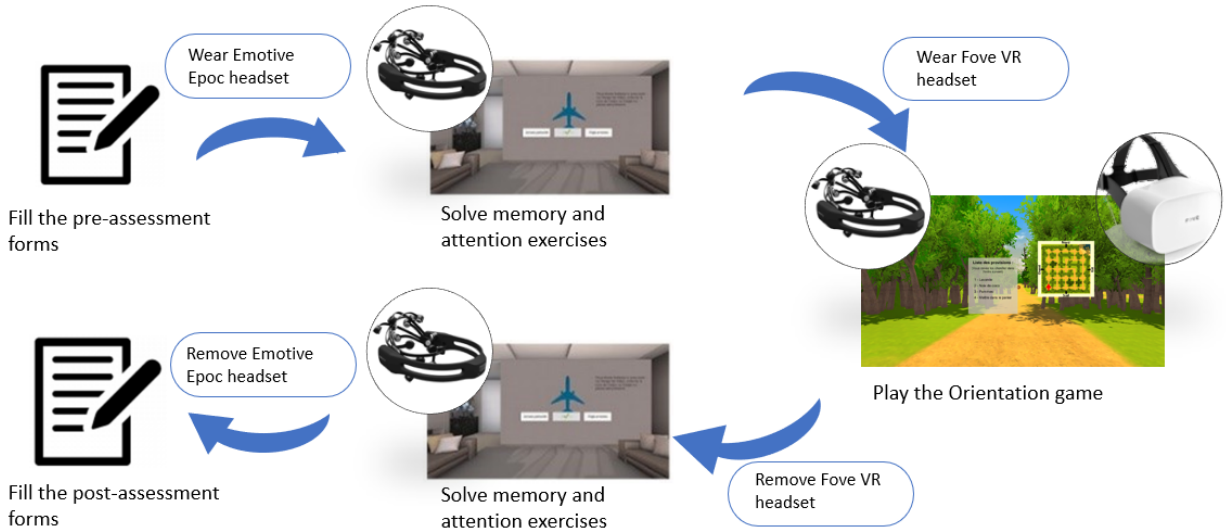


Figure 3.7. Process of the experiments

EEG headset was removed, and the participants were asked to fill up the following post-experiment forms: PANAS scale, cyber-sickness, the AttrackDiff 2 [**51**] and a self-report form. Figure 3.7 summarises the different steps of the process of the experiment.

3.4. Data Collection

The pre-assessment forms filled by participants contribute as the demographic information which can be used for analysis after masking. After the participants put on the Emotive Epoc at the start of experiments, the emotion values are recorded during the entire training session. The EEG device transmits brain signals and emotion values at the rate of up to 128 samples per second. However, to reduce noise, to reduce high fluctuation in recorded value and to simplify the calculations, the emotion values are smoothed to provide one sample per second. During the VR training session, the location of participants in the environment, the emotion values, the hint provided by the guidance system and the cause of activation of hint is recorded every second.

Chapter 4

Results and Discussion

The primary objective of our research was to study if it is possible to generalize the effects of cognitive training in a spatial orientation tasks to memory and attention in people with cognitive decline. In this study, we defined our scope to study the immediate after-effects of playing the orientation game. The second objective of this research was to analyze the effect of hints provided by the guidance system on emotions of the participants. Some of the results discussed below were presented in 16th International Conference on Intelligent Tutoring Systems'. The accepted paper is presented in appendix A.

4.1. Performance in Attention and Memory Exercises

In the first attention exercise, the participants have to repeat a numerical sequence after hearing it once. Each sequence in the exercise is scored for the number of digits a participant gets right. Thus the first sequence is scored out of five, and the maximum score in the second sequence is three. Next, in the second attention exercise, the participant receives a positive score for each time they identify the letter 'A' in the sound sequence out of a total of eleven times. Further, in the third attention exercise, participants have to identify six objects by their initial letter. Thus, the maximum score in this exercise is six. Among the memory exercise, in the first memory exercise, the participants are examined on six items that they have to memorize and scored out of six. The second memory exercise is tested on two levels of difficulties with a total of ten circles which is the maximum score of this exercise. For the last memory exercise, two examples of working memory are required to be solved and a correct answer in each gives one point each. Hence, the maximum score in the last memory exercise is two. The scores obtained by participants in each experiment are normalized out of ten and further analysis is performed.

	Percentage Improvement in Attention Exercises			Percentage Improvement in Memory Exercises		
	1	**2**	**3**	**1**	**2**	**3**
P1	0.0	18.2	16.7	33.3	0.0	50.0
P2	0.0	0.0	0.0	0.0	0.0	0.0
P3	12.5	0.0	0.0	66.7	30.0	50.0
P4	0.0	0.0	0.0	0.0	-10.0	50.0
P5	25.0	0.0	0.0	16.7	-30.0	50.0
P6	0.0	0.0	0.0	0.0	60.0	50.0
P7	12.5	-27.3	0.0	-33.3	-20.0	50.0
P8	0.0	0.0	-16.7	-16.7	0.0	0.0
P9	0.0	0.0	0.0	0.0	30.0	0.0
P10	25.0	0.0	0.0	0.0	30.0	50.0
P11	25.0	18.3	0.0	0.0	0.0	0.0
P12	0.0	0.0	0.0	0.0	0.0	0.0
P13	0.0	0.0	0.0	16.7	0.0	0.0
P14	0.0	0.0	0.0	0.0	30.0	50.0
P15	0.0	0.0	0.0	33.3	60.0	0.0

Table 4.1. Comparing performance improvement in the attention and memory exercises before and after playing the orientation game.

First, we compare the performance in the attention and memory exercises conducted before and after the orientation game. The change in score for all the participants are reported in table 4.1. We can note that in attention exercise 1, five participants showed an improvement, while the rest maintained the same number of correct answers. In exercise 2, two participants scored better after playing the orientation game, P7 scored lesser, and the rest kept the same level of performance. Finally, in the third exercise, P1 showed an increase, while P8 had a decrease, and the rest had the same score before and after playing the orientation game. In the first exercise, the mean improvement was 7% for all the participants. In the second exercise, there was negligible improvement and no change in average performance was observed in the third exercise.

Next, we analyze the performance of the participants in the memory exercises. We notice that in the fourth exercise, which is a test of contextual memory, three participants showed an increase in the number of correct answers, while four of them displayed a decrease, and there was no change in the rest of the participants. The mean of the increment in score for this exercise was 1% among all the participants. In the second memory exercise, which is a test for short-term memory, six participants improved their score, while three had a decrease in performance, and there was no change in scores for the rest of the participants. The average change in the second memory exercise was 12.00%. Finally, in the last memory exercise aimed at screening the working memory, the mean increase in score was 27%

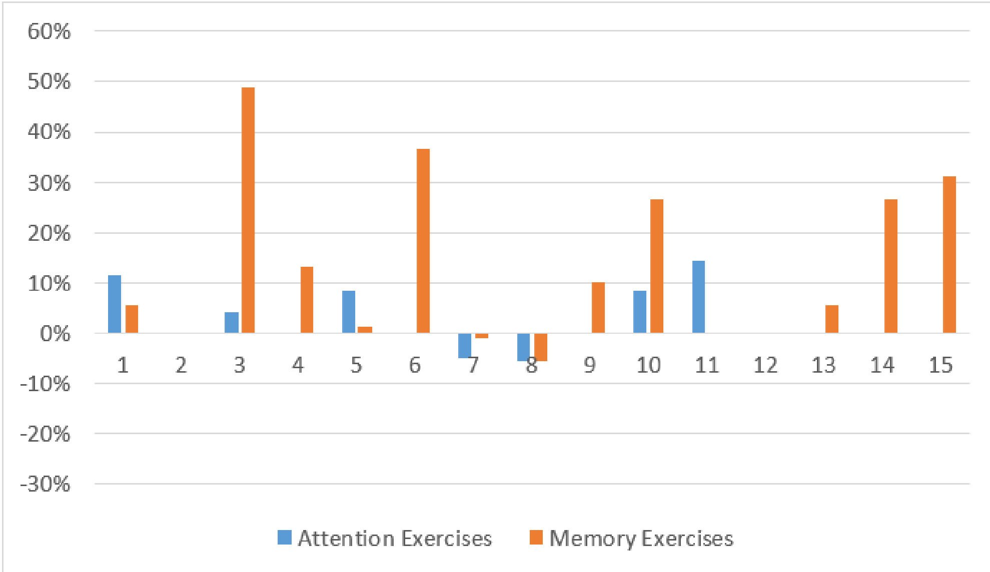


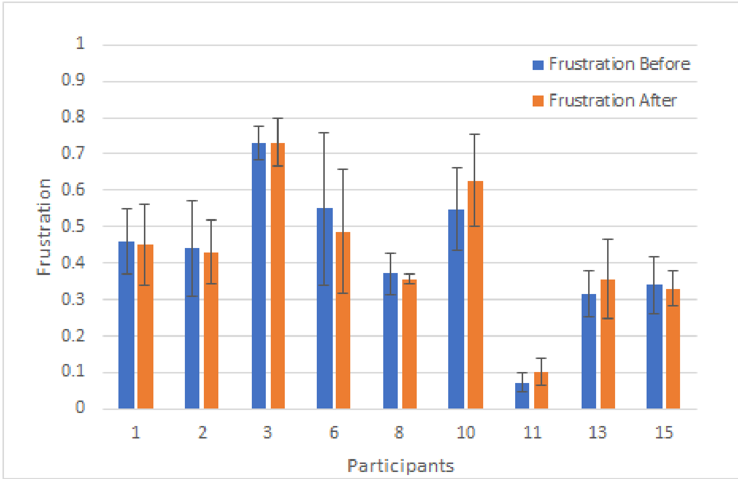
Figure 4.1. Histogram of average performance improvement in attention vs cognitive exercises for all the participants

with eight participants reporting better scores while the rest retained the same level of performance (none of them reported a decrease of performance).

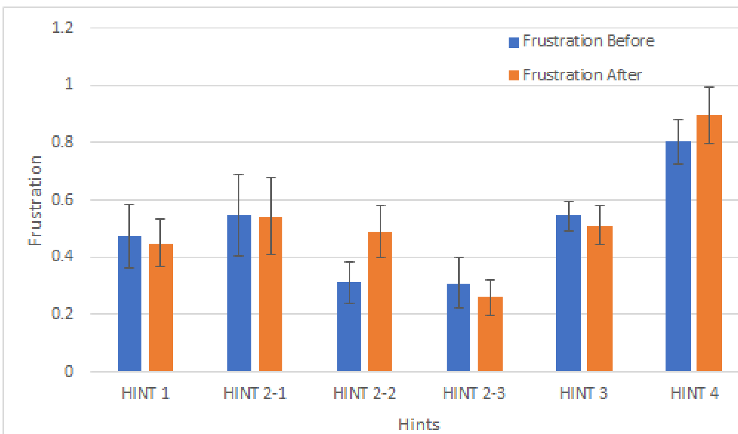
These results show a noticeable increase in memory exercise when compared to attention exercise. However, different participants showed different levels of change and some enhanced their scores more than others. For example, P1 improved his/her score in a total of four exercises (two attention and two memory exercises). Similarly, P3 improved his/her score in four exercises(one attention, two memory exercises). P2 and P12 scored the same scores in each of the exercises, while P7 reported the highest drop in performance with lesser scores in one attention and two memory exercises. These results lend some support to the hypothesis that orientation based cognitive training tasks may have positive effects that transfer to short-term memory and working memory. Yet, how well these results generalize needs to be examined by allowing the participants to train regularly over a longer period of time.

4.1.1. Effect of Guidance System on Emotions

To understand if the hints are effective in reducing frustration and maintaining engagement in the cognitive training, we analyzed the emotions of the participants before and after the hints are provided. The guidance system provided at least one or more hints to 9 out of the 15 participants. Participants P8, P10, P11 and P13 were provided one, two, one and one hint respectively, while the rest of the participants received at least five or more hints. In the rest of the cases, participants did not have any difficulty in completing the quests. Change



(a) Frustration values of each participant for all the hints, reported as the average taken 10 seconds before and after the hints were provided.



(b) Frustration values for each hint type 10 seconds before and after the hints were provided, averaged over all the participants who received hints.

Figure 4.2. Effect of hints on frustration. The vertical line in each column shows the range within one standard deviation.

in emotions of the rest of the participants did not trigger the hints from the guidance system.

In figure 4.2(a) we notice that in six out of nine participants, there was a decline in average frustration in next ten seconds after the hints were provided to them. However, for the participants P10, P11 and P13, average frustration increased slightly after the hints. For the different types of hints, we can notice in figure 4.2(b) that except for hint level 2-2 and hint level 4, the average values of frustration for all the participants were lesser in the next ten seconds after the hints were provided.

We can notice that hint level 2-2 that provides a warning message: 'You're too far. Try to take few steps back.', is not in a positive mood. This led the participants to believe that they might be doing something wrong leading to higher frustration. Hint level 4 displays the complete path to the item's location. But since we configured the hint to appear for only four seconds, this time may not enough for the participants to memorize the complete path to the item's location, which may have led to a higher level of frustration.

Though it may seem that the decrease in some cases is not significant, the stabilization of frustration after a continuous increase before hint is given shows that it is helpful. In the after-experiment survey, six out of nine participants who were given hints, found the hints to be helpful to solve the quests, which further supports the utility of the real-time guidance system in 3D VR orientation games.

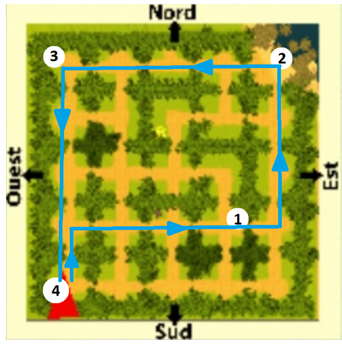
Chapter 5

Improving Guidance System using Reinforcement Learning

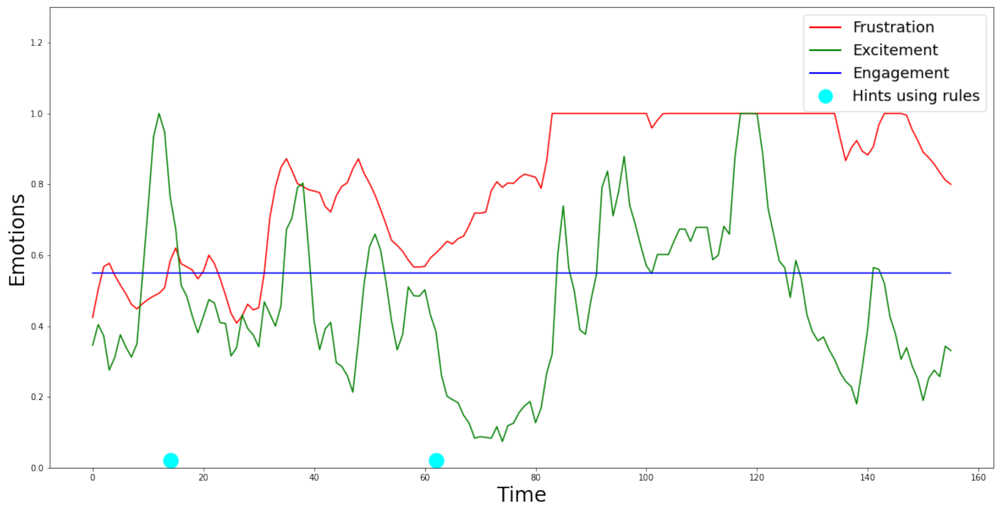
5.1. Limitations of Rule-based Guidance System

We analyzed that rule based guidance system was effective in helping people in completing objectives of the orientation games. However, the rule-based system was based on a set of assumptions and empirically decided hyper-parameters which were not always applicable. To go into more detail, we take an example of hints provided to the participant P10 as shown in figure 5.1.

Figure 5.1(a) shows the movement of the participant in the VR environment. Initially a rapid drop in excitement exhibited by the participant triggers hint level 1 at 17 seconds in the game, which is shown by a blue dot at the X-axis in figure 5.1(b). P10 reached the location of lavenders, the first item to be collected at 27 seconds. But it was not for another 13 seconds that he/she was able to find the lavender plants by turning around in the environment. At this point, we observe a sharp increase in the frustration which is evident in the graph around 30 seconds. This behavior was noted in some other participants as well where they could comfortably reach the location of the item but could not spot the item easily by looking around. Since hints were not provided when the person had already reached the target location, no hint was displayed in this case. This revealed a drawback in the assumption in the rule-based system, that once the player reaches the target location, he/she does not need a hint to further help them locate the item. This rise in frustration on reaching the first item was witnessed in many participants and was caused due to the unfamiliarity with such training programs and in some cases difficulty in orienting themselves within the environment. We notice a drop in frustration when the participant collects the lavenders at 41 seconds, followed by slight increases when the instruction for the second item was displayed. After the instructions, the participant had no issues in reaching



(a) Path taken



(b) Changes in emotions, and hints provided by rule based guidance system

Figure 5.1. Path taken by P10 and the rule based hints provided.

the location of the coconuts, the second item to be collected, in another twelve seconds. At this time, we again notice a gradual increase in the frustration around 63 seconds until the item was collected at 80-second mark. Once again, since the participant was already at the target location, no hint was displayed. The frustration momentarily dropped around the 80-second timestamp, after the item was collected, but stayed near the highest value of 1.0 for the rest of the game. The frustration slightly dropped again at 130 seconds when apples were collected, which were the third item in the list. Finally, the participant reached the basket to drop all three items at the 155-seconds mark that also marks the end of the game. We notice that no hint was displayed during the rest of the training. The rule-based system checks for any criteria which satisfy the condition to give a hint after an interval of 15 seconds. The person was moving in the right direction and did not stop at one location for more than 10 seconds, hence the hints due to '*no movement*' or '*away from target*' were not applicable. Further, hints due to emotions were not activated as whenever the criteria for hint was checked, the net '*emotiscore*' of the 'excitement' was always greater than frustration.

This analysis underscores some fundamental limitations of the rule-based guidance system. The fifteen-second time interval to check for hint condition was suitable for some participants, but it was not good in the case of some others such as P10 in this matter. We also observed that the participant faced high frustration even when he/she could reach the item location. In such cases, a hint saying that the item can be found by turning around in the environment could have been helpful. Another take away in the above example was that a high level of frustrations is more important to consider compared to the excitement.

Though one option to improve the guidance system was to update the rules by taking into account the data gathered from the experiments, it is difficult to come up with a set of rules which could generalize across all participants in all scenarios. A better way would be to design a system that could learn to provide an optimal number of hints and at a suitable time so as to reduce the frustration of the participants. Once, the system recognizes a proper time to present hint, the next step would be to identify the level of hints to be given which can be guided by a simpler set of rules. Therefore, in this case, we implemented a reinforcement learning(RL) agent that would identify the correct time to give a hint by analyzing the player's frustration, after which the rule-based system would simply decide the type of hint to be presented based on the scenario.

5.2. Reinforcement Learning: Background

Reinforcement learning(RL) is learning the actions to take so as to maximize numerical rewards [**1**]. It is the intuitive way of learning in humans where we learn from the outcomes of our actions and in turn take actions so as to get best outcome. We state the Markov decision processes(MDP) which is the formal problem setting of reinforcement learning.

5.2.1. Markov Decision Process

Reinforcement learning problems can be framed as an optimization problem in an incompletely-known Markov decision process. Formally the decision taker is called the agent and the medium it interacts with it called an environment. An MDP is a 5-tuple: $\langle \mathcal{S}, \mathcal{A}, \mathcal{P}, \mathcal{R}, \gamma \rangle$, where $\mathcal{S}$ is the set of all possible states; $\mathcal{A}$ is set of possible actions which the agent can take; $\mathcal{P}$ is the state transition probability; $\mathcal{R}(S_t, A, S_{t+1})$ is the reward function which gives a numerical feedback to the agent; and γ is the discount factor in the range $0 \leq \gamma < 1$, which specifies the relative weight given to immediate rewards relative to the rewards in future.

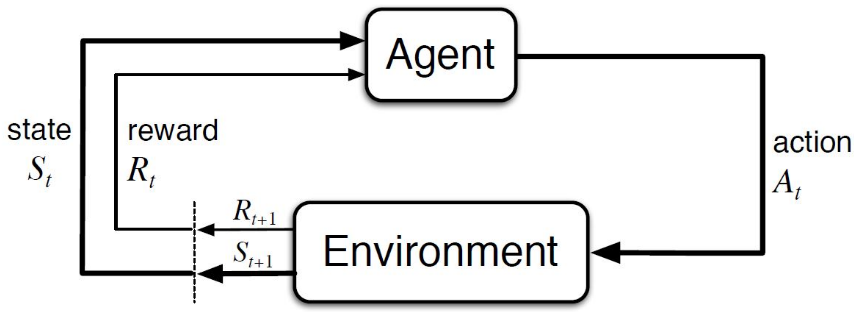


Figure 5.2. The agent-environment interaction in a Markov decision process. Section 3.1 Sutton and Barto [**1**]

In an MDP, the agent selects actions from the set of actions, and the environment gives a feedback in the form of a numerical reward and present new state to the agent which depends on the action taken. At each time step t, the agent finds itself in a state $S_t \in \mathcal{S}$. Selecting a new action $A_t \in \mathcal{A}$ from this state takes agent to a new state $S_{t+1} \in \mathcal{S}$ of the environment and provides a numerical reward $R_{t+1} \in \mathcal{R}$ in the next time step. The agent and the environment react repeatedly leading to a sequence of state, action and rewards as following:

$$S_0, A_0, R_1, S_1, A_1, R_2, S_2, A_2, R_3, ... \tag{5.2.1}$$

In an MDP, all the states have Markov property, i.e. the next state is conditionally independent of all previous states and actions.

$$\mathbb{P}[S_{t+1}|S_t] = \mathbb{P}[S_{t+1}|S_1, \ldots, S_t] \tag{5.2.2}$$

Every MDP has a state transition probability function $\mathcal{P}(S_{t+1}|S_t, A_t)$ which defines the probability of the transition of the environment from state S_t to the next state S_{t+1} on taking an action A_t. The probabilities given by $\mathcal{P}$ define the characteristics of the environment.

In a sequence of transitions, we expect to maximize the expected return from current state. In simplest case, the return is sum of discounted reward obtained in future.

$$G_t \doteq R_{t+1} + \gamma R_{t+2} + \cdots = \sum_{k=0}^{\infty} \gamma^k R_{t+k+1}, \tag{5.2.3}$$

where $\gamma \in [0, 1]$ is the discount factor. The discount factor gives higher weight to immediate reward and penalizes the rewards received in future by a factor γ^{k-1} owing to the uncertainty in the future.

The agent does not have prior information about either $\mathcal{P}(S_{t+1}|S_t, A_t)$ or $\mathcal{R}(S_t,A,S_{t+1})$ and must learn by taking different actions and receiving feedback. The family of Rl algorithms which learn from experience and do not rely on knowing the behaviour of environment are known as *model-free* algorithms. Model-free algorithms are further subdivided into on-policy and off-policy. In on-policy learning, the value function is learned from action taken as per the current policy $\pi(a|s)$, while in off-policy learning, value function is learned by taking different actions and updating value function greedily based on the best action so far.

5.2.2. Policy and Value Functions

Policy guides the behaviour of the agent in an unknown environment. A stochastic *policy* $\pi(a|s)$ is a probability function which gives the mapping from state $S_t = s$ to the probability of selecting action $A_t = a$ from that state. In reinforcement learning, the problem consists of finding the optimal policy for the agent so as to maximize the expected discounted return G_t .

Value functions give a quantitative measure of how good a state or action is for the given policy π, in terms of the future rewards received from that state or on taking that action. The *state-value function* is defined as the expected return from given state $S_t = s$.

$$V_\pi(s) = \mathbb{E}_\pi[G_t|S_t = s] \tag{5.2.4}$$

Similarly, the *action-value function* is defined as the expected return on taking an action $A_t = a$ from a given state $S_t = s$.

$$Q_\pi(s, a) = \mathbb{E}_\pi[G_t|S_t = s, A_t = a] \tag{5.2.5}$$

There is always an *optimal policy* π_* which is equal or better than all the other policies. Value functions of optimal policy produce the maximum return.

$$V_{\pi_*}(s) = V_*(s) = \max_\pi V_\pi(s) \tag{5.2.6}$$

$$Q_{\pi_*}(s, a) = Q_*(s, a) = \max_\pi Q_\pi(s, a) \tag{5.2.7}$$

5.2.3. Bellman Equations

Bellman equations define the recursive relationship of value function as the sum of the immediate reward plus the discounted future values. For state-value functions:

$$\begin{aligned} V(s) &= \mathbb{E}[G_t|S_t = s] \\ &= \mathbb{E}[R_{t+1} + \gamma R_{t+2} + \gamma^2 R_{t+3} + \ldots |S_t = s] \\ &= \mathbb{E}[R_{t+1} + \gamma G_{t+1}|S_t = s] \\ &= \mathbb{E}[R_{t+1} + \gamma V(S_{t+1})|S_t = s] \end{aligned} \tag{5.2.8}$$

Similarly for action-value function:

$$
\begin{aligned}
Q(s,a) &= \mathbb{E}[R_{t+1} + \gamma V(S_{t+1}) \mid S_t = s, A_t = a] \\
&= \mathbb{E}[R_{t+1} + \gamma \mathbb{E}_{a\sim\pi} Q(S_{t+1}, a) \mid S_t = s, A_t = a]
\end{aligned}
\tag{5.2.9}
$$

Both state-value function and action-value function can be expressed in terms of policy and transition probability (P) as follows.

$$
V_\pi(s) = \sum_{a\in\mathcal{A}} \pi(a|s)\Big(R(s,a) + \gamma \sum_{s'\in\mathcal{S}} P^a_{ss'} V_\pi(s')\Big) \tag{5.2.10}
$$

$$
Q_\pi(s,a) = R(s,a) + \gamma \sum_{s'\in\mathcal{S}} P^a_{ss'} \sum_{a'\in\mathcal{A}} \pi(a'|s') Q_\pi(s',a') \tag{5.2.11}
$$

Optimal state-value and action-value given in equation 5.1.6 and 5.1.7 respectively are received on following the optimal policy π_*. In optimal case, the Bellman equations of the value functions can be written as:

$$
V_\pi(s) = \sum_{a\in\mathcal{A}} \pi(a|s)\Big(R(s,a) + \gamma \sum_{s'\in\mathcal{S}} P^a_{ss'} V_\pi(s')\Big) \tag{5.2.12}
$$

$$
Q_*(s,a) = R(s,a) + \gamma \sum_{s'\in\mathcal{S}} P^a_{ss'} \max_{a'\in\mathcal{A}} Q_*(s',a') \tag{5.2.13}
$$

5.2.4. Q-Learning Algorithm

Q- learning introduced by Watkins and Dayan [**52**] is an off-policy model-free algorithm to find the optimal policy which maximizes the expected value of the total reward obtained in all future steps, starting from the current step. Algorithm 5.2.1 shows the Q-learning algorithm as it appears in the RL textbook by Sutton and Barto [**1**].

Algorithm 5.2.1. Q-learning for estimating optimal policy π_*.

Algorithm parameters: step size $\alpha \in (0,1]$, small $\epsilon > 0$
Initialize $Q(s,a)$, for all $s \in \mathcal{S}, a \in \mathcal{A}$, arbitrarily except that $Q(terminal, .) = 0$
For each episode:
  Initialize S_t
  For each step of episode:
    Choose A_t from S_t using policy derived from ϵ-greedy Q
    Take action A_t, observe R_t, next state S_{t+1}
    $Q(S_t, A_t) \leftarrow Q(S_t, A_t) + \alpha(R_{t+1} + \gamma \max_{a\in\mathcal{A}} Q(S_{t+1}, a) - Q(S_t, A_t))$
    $S_t \leftarrow S_{t+1}$
  until S_t is terminal

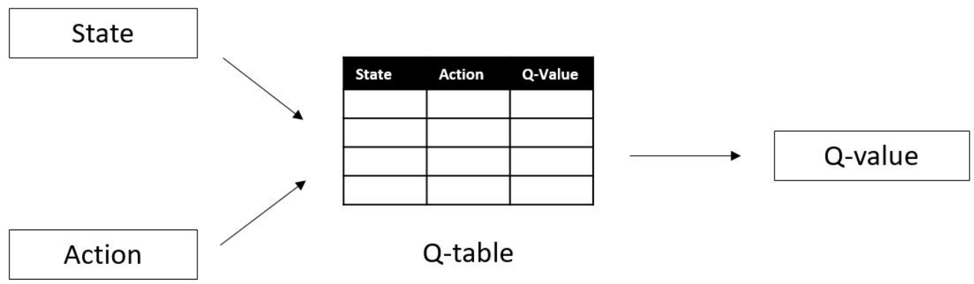


Figure 5.3. Q-Learning

At each time step, we start from state S_t and pick action according to ϵ-greedy Q values such that $A_t = \arg\max_{a \in \mathcal{A}} Q(S_t, a)$. An ϵ-greedy approach to select actions ensure that we explore non-greedy actions with probability ϵ and avoid getting stuck in local maximum. On taking action A_t, the environment gives a reward R_{t+1} and transitions into the next state S_{t+1}. Next, we update the action-value function:

$$Q(S_t, A_t) \leftarrow Q(S_t, A_t) + \alpha(R_{t+1} + \gamma \max_{a \in \mathcal{A}} Q(S_{t+1}, a) - Q(S_t, A_t)) \tag{5.2.14}$$

This process is repeated for a given number of epochs or till the Q-values converge to a value in acceptable range. Q-value of each state-action pair is updated in a table called Q-table as shown in figure 5.3. Once the action-values for each state-action pair are updated, the agent can simply take the action which give best possible return from that particular state.

5.2.5. Deep Q-Learning Networks

Learning Q-values for each state-action pair becomes infeasible and computationally expensive when the state is continuous or arbitrarily large. In such cases, the goal changes to learning near-optimal approximation of action-value function given by equation 6.1.13. A neural network is most commonly used function approximator for this purpose. In Deep Q-Learning Networks(DQN), two neural networks with same architecture are used: first network is used as the funtion approximator for action value-function, and second network with frozen parameters to update target Q-values. After every C iterations, the parameters from the first network are copied to the target network. A technique known as experience replay is used to remove correlations in the observation sequence and smoothing over changes in the data distribution. Algorithm 5.2.2 explains the training of deep q-learning networks as it appeared in the work by Volodomyr et al. [**53**].

The agent's experience at each time step, $e_t = (s_t, a_t, r_t, s_{t+1})$ is stored in a data set, $D = e_1, e_2, ..., e_n$ into a replay memory. First an action is selected using $\epsilon - greedy$ policy. This ensures that we select a random action a with probability ϵ, and select the action with maximum Q-value ($a = argmax(Q(s,a,\theta))$) with probability $1 - \epsilon$. The selected action is performed and the transition (s, a, r, s') is stored in the replay memory D. Next, a sample of random mini-batches of transition ϕ is selected at random from the replay memory D to calculate the loss:

$$Loss = (R_j + \gamma max_{a'} Q(\phi_{j+1}, a'; \theta) - Q(\phi_j, a; \theta))^2 \tag{5.2.15}$$

Finally, gradient descent with respect to the network parameters θ is performed to minimize the loss. After every C iterations function approximator network weights are assigned to target network. Figure 5.4 demonstrates the input and the output of the deep Q-learning networks.

Algorithm 5.2.2. Deep Q-learning with Experience Replay.

Initialize replay memory D to capacity N
Initialize action-value function Q with random weights θ
Initialize target action-value function $\hat{Q}$ weights $\theta' = \theta$
For each episode:
    Initialise sequence $s_1 = x_1$ and pre-processed sequenced $\phi = \phi(s_1)$
    For each step of episode:
        Choose A_t from S_t using policy derived from ϵ-greedy Q
        Take action A_t, observe R_t, next state S_{t+1}
        Set $S_{t+1} = S_t, A_t, x_{t+1}$ and pre-process $\phi_{t+1} = \phi(S_{t+1})$
        Store transition $(\phi_t, A_t, R_t, \phi_{t+1})$ in D
        Sample random mini-batch of transitions $(\phi_j, A_j, R_j, \phi_{j+1})$ from D
        If terminal ϕ_{j+1}
            Set $y_j = R_j$
        If non-terminal ϕ_{j+1}
            Set $y_j = R_j + \gamma max_{a'} Q(\phi_{j+1}, a'; \theta)$
        Perform a gradient descent step on $(y_j - Q(\phi_j, a; \theta))^2$ with respect to network parameters θ
        Every C steps, reset $\hat{Q} = Q$
    until S_t is terminal

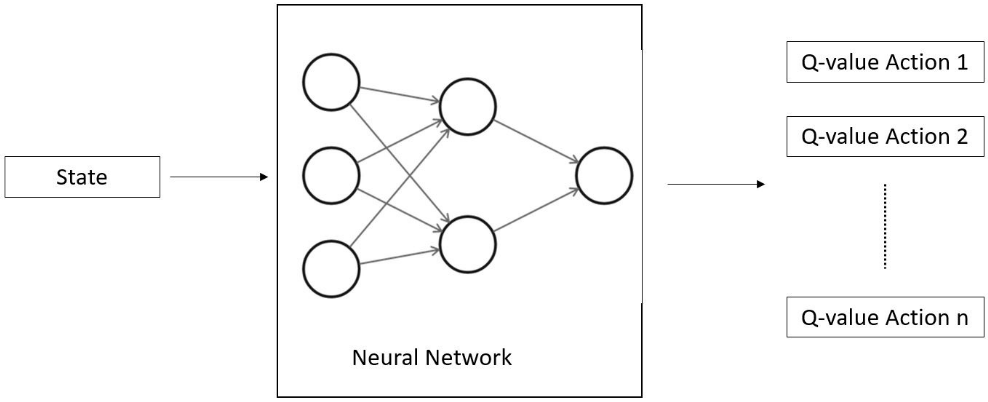


Figure 5.4. Learning Q-values with Deep Q-learning Network

5.3. Guidance System in Orientation Game as Reinforcement Learning Problem

The purpose of the guidance system is to effectively help the participants so that they are able to complete the objectives of the orientation game without being frustrated to the extend that they have the tendency to give up before completing the training program. However, the hints must not be provided to the participants throughout the game, as it will make the tasks less challenging and lead to minimum learning. The guidance system can be considered as the agent which must learn the optimized way of providing hints so that the participants are able to complete tasks within acceptable level of frustration using minimum number of hints in the process. We should not that in context of the RL, the environment consists of human carrying the tasks and the VR environment, both of which are external to the agent providing hints,

5.3.1. State, Actions and Rewards

The first task is to identify the states and actions in the guidance system. While visually the the VR environment is in the form of 5x5 maze, the maze in this case is represented by a 10x10 2D array, in which paths are marked by 1 and inaccessible areas formed by trees are marked by 0. The location of the person solving tasks in the garden is marked by 0.5 while the location of items is marked by 0.75. Next, we include emotions as the state of the person at the current moment. The change in the emotion values give an idea about the effect of several in-game elements on the person since last time step. The number of items remaining to be collected is included as it may affect the movement of the person in the VR. The last boolean value checks if a hint is on display so that agent learns to not give another hint while previous hint is being presented. The complete state of the environment external to the agent is given by:

$$State, s_t = [M_t, [E_t^{[f]}, E_t^{[en]}, E_t^{[ex]}], [\Delta E_t^{[f]}, \Delta E_t^{[en]}, \Delta E_t^{[ex]}], (N - C_t), is\_hint\_on] \qquad (5.3.1)$$

where,
M_t = 2-D array representing state of the maze environment
$[E_t^{[f]}$ = Frustration of the person at time t
$\Delta E_t^{[en]}$ = Engagement of the person at time t
$\Delta E_t^{[ex]}$ = Excitement of the person at time t
$\Delta E_t^{[f]}$ = Change in frustration since last time-step
$\Delta E_t^{[ex]}$ = Change in excitement since last time-step
$\Delta E_t^{[en]}$ = Change in engagement since last time-step
N = number of items to be collected (4 in this case)
C_t = number of items collected at time t
is_hint_on = boolean value to express if a hint is being presented

It was observed in the experiments with rule-based hints that in most cases hint level 1 and 2 were displayed. Hence, to simplify the training, we limited the task of the agent to identify if a hint needs to be presented or not. Thus, the action-space consists of two actions: 1) show a hint, 2) do nothing.

We can safely assume that high amount of time spent in completing the tasks is undesirable since the participants who found the tasks easier spent lesser time in completing all the objectives. The time spent in the VR environment can be interpreted as negative feedback to the agent at every time step. For every item collected by the human, a positive reward is given, and when the human completes the game by collecting all items and returns to the start position, a high positive reward is given. Since we want to restrict a high level of frustration, a negative reward is provided if the human exhibits high frustration values sustained for an interval greater than ten seconds. A high negative reward is given if the agent furnishes a hint when another hint is on display. We added a negative reward to penalize the hints if they were given too often, which would decrease every second a hint was not furnished and resets whenever a hint is given.

5.3.2. Generating Simulated Episodes

To simulate the movement of a person in the VR environment, it was assumed that the person moves towards the target with probability p, while she/he moves in any available direction randomly with probability $1 - p$. Under the influence of a hint, this probability would increase by a factor k, and the influence wears off with every passing i seconds by a factor of $1/i$. The hyper-parameters p and k were set to 0.6 and 1.25 respectively. The mean and standard deviation of the normal distribution for each of the emotions - frustration, excitement and engagement - was obtained from the experiments previously conducted with rule-based hints as shown in table 5.1 . To simulate each of the emotion, a normally distributed random values corresponding to mean and standard deviation of each emotion was generated for each timestamp. A gaussian noise with standard deviation 0.05 was added to each of the generated emotion value for regularization effect. Any value thus obtained below 0 was clipped to 0, and values higher than 1 were clipped to 1.

	Mean	SD
Frustration	0.523	0.241
Excitement	0.397	0.240
Engagement	0.642	0.078

Table 5.1. Mean and standard deviation of emotion values of fifteen participants during orientation game.

5.3.3. Training

Deep Q-learning network with experience replay is ideal for episodic tasks[1] with discrete action space. Since, emotions values are continuous variable, and the action space was discrete in our case, DQN with experience replay (algorithm 5.2.2) was the choice of algorithm for training the reinforcement learning model for the agent. Two network with parameters θ and θ' were initialised with same network architecture. The input to the network was a tensor with the nine state variables mentioned in equation 6.2.1. The network consisted of two hidden layers, both of which had hundred nodes, equal to the size of the 2D-array representing VR environment. The output layer size was two which is the same as the number of actions, to return q-values for each action which are either to show a hint or do nothing. The network was trained using 'adam' optimizer with minibatches of size 128.

5.3.4. Evaluation

The RL agent was evaluated on the data of the fifteen participants to check if it was able to identify the time to provide hints better than the rule-based guidance system. To evaluate the learned agent, we converted the movement and emotions data of the human into the input state as in equation 6.2.1 and output the greedy actions at each time which

[1]Episodic tasks start from a particular state and have a terminal state.

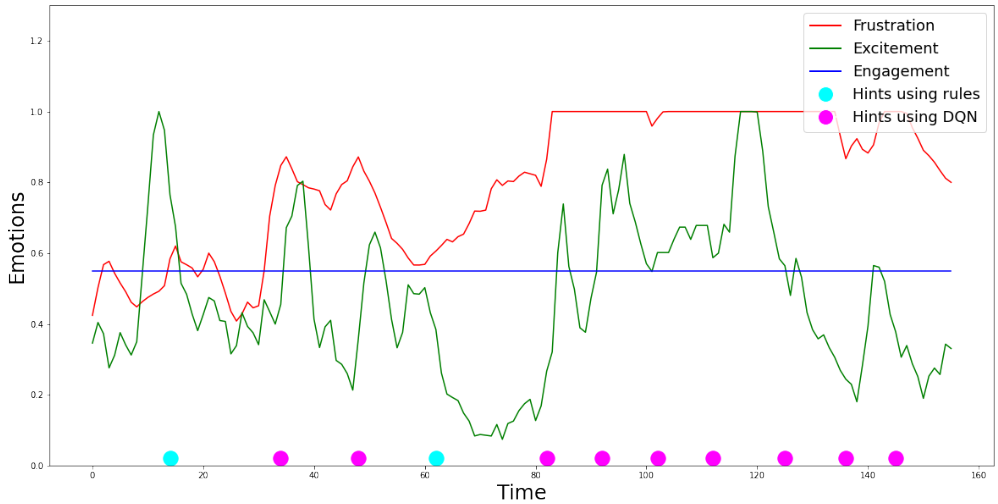


Figure 5.5. Hints provided by rule-based system compared to hints provided by RL agent for P10

corresponds to the maximum q-values in that state.

As an example, we again take the participant P10. Figure 5.5 presents a comparison between the timestamps at which hints were provided by the rule-based system vs the timestamps at which hints are recommended by the RL agent. The blue dots on the X-axis are the timestamp at which hints were provided by the rule based guidance system in actual experiments, while the pink dots represent the timestamp at which hints are suggested by the RL agent. It can be noted that hints provided by the latter correspond to either an increase in frustration or high value of frustration in general. It implies that the agent learned to minimize the negative rewards given when there is a rise in frustration, irrespective of the excitement levels. There is some time gap between successive hints which is a consequence of penalizing the reward earned by the agent if given in a short time interval. The number of hints given in this case is more than the number of hints provided by the rule-based system. Appendix C shows a similar comparison between the hints provided by the two systems for all the participants. We can notice that in most of the participants, RL based system proves more sensitive to increase in frustration and provides the hints more frequently compared to rule-based guidance system.

Chapter 6

Conclusion and Future Work

In this research, we set out to study the effect on attention and memory abilities in older adults, after undergoing a cognitive training program that required spatial orientation skills. To this end, we implemented a 3D virtual reality orientation game using Unity 3D and examined the improvement in attention and memory exercises after playing the game. In the attention exercises, most of the participants kept similar scores, and no significant increase or decrease in performance was observed. However, based on the performance in memory exercises, we can conclude that the orientation game helped in enhancing short-term memory and working memory skills. This conclusion promotes the idea of near transfer [**6**] that training a particular ability can transfer to tasks that use the same components of the human brain. In this case, the hippocampus is the part of the brain which is responsible for spatial orientation that enables navigation and long-term memory formation by consolidating short-term memory.

Furthermore, we created an efficient guidance system to help reduce frustration during the training task. The guidance system is responsible to provides hints in cognitive training tasks by utilizing the emotions of the participants. Initially, we created a rule-based system that used the participant's movement pattern and the recent changes in her/his emotions to identify befitting situations to provide hints. We observed that most of the participants benefited from this guidance system and found it to have a positive impact on their game-play. However, hint level 2-3 caused an increase in frustration of the participants, owing to the cautious tone of this particular hint. The hint level 4 showing the complete path within the map was an example of a hint which was difficult to follow, as the participant had to memorize the complete path in four seconds. This also led to a rise in the frustration of participants. Thus, we can conclude that while hints are helpful to older adults in cognitive tasks, given that the messages in hints keep a positive feel and are easy to understand.

Although the guidance system was useful, the rule-based hints suffered from some limitations due to high variation in emotions of the participants and assumptions in that respect. Hence, to improve the guidance system, we used a reinforcement learning agent to identify the optimal situations to recommend hints in the orientation game. The deep Q-learning network with experience replay is ideal for this scenario as this approach is sample efficient and works in problems with discrete action space. In our case, we used DQN to identify preferred scenarios to provide hints. The network was trained using simulated data generated based on the data gathered during experiments conducted by using the rule-based guidance system. By comparing the recommended time to deliver hint by the RL agent with the rule-based hint, we notice that RL agent adapts significantly better using the emotions.

6.1. Future Work

It is a real challenge to create a cognitive training program that can suit the needs of every individual. Different people react to the same situations differently. In our experiments, although the same protocol of memory and attention exercise followed by the VR Orientation game was administered to all the participants, each of them manifested different levels of frustration, engagement, and excitement. Some participants experienced high frustration while playing the game, while others displayed little frustration and high engagement. Some participants were able to complete the tasks faster than others, and some of them required more hints than others. Creating a generalized training program accommodating such levels of variation would be tedious. A comparatively modest challenge would be to personalize the training program for each individual on account of their interaction with the environment and performance in the game.

In our future work, we intend to change the level of difficulty of tasks dynamically based on the player's emotions and the time taken to complete the objectives. The emotions of the participant can be used to adjust the difficulty level of the orientation game in real-time. The next object can be placed at a more complicated location in the maze to make the game more difficult. We also aim to change the number of items in the game conditioned on the current difficulty level and the participant's performance. More advanced features would be to add fog and other elements to limit the sight. Feedback based on emotions can be used in a variety of training exercises to adapt the program to suit every individual. We intend to explore this in other VR training programs.

To sum up, it's a long way to go before we arrive at affordable interventions that can be administered at home for slowing the cognitive decline in older adults. It is more beneficial to have CT programs that can extend the benefits of learning to daily life tasks. Training programs that enhance overall mental health may further reduce the risks of emotional instability in afflicted individuals. Cognitive training using tasks in VR environments offer some promise of being helpful as an affordable solution in the coming times. The aspiration of the cognitive training tasks which can be undertaken at home with minimum supervision could soon be a reality with more sensitive EEG devices, affordable VR headsets, better emotion recognition models, and more adaptable virtual environments.

References

[1] R. S. Sutton and A. G. Barto, *Reinforcement learning: An introduction.* MIT press, 2018.

[2] A. Association, “2019 alzheimer’s disease facts and figures,” *Alzheimer’s & Dementia*, vol. 15, no. 3, pp. 321–387, 2019.

[3] A. Imfeld, M. S. Oechslin, M. Meyer, T. Loenneker, and L. Jancke, “White matter plasticity in the corticospinal tract of musicians: a diffusion tensor imaging study,” *Neuroimage*, vol. 46, no. 3, pp. 600–607, 2009.

[4] E. A. Maguire, D. G. Gadian, I. S. Johnsrude, C. D. Good, J. Ashburner, R. S. Frackowiak, and C. D. Frith, “Navigation-related structural change in the hippocampi of taxi drivers,” *Proceedings of the National Academy of Sciences*, vol. 97, no. 8, pp. 4398–4403, 2000.

[5] P. A. Reuter-Lorenz and K. A. Cappell, “Neurocognitive aging and the compensation hypothesis,” *Current directions in psychological science*, vol. 17, no. 3, pp. 177–182, 2008.

[6] E. Dahlin, A. S. Neely, A. Larsson, L. Bäckman, and L. Nyberg, “Transfer of learning after updating training mediated by the striatum,” *Science*, vol. 320, no. 5882, pp. 1510–1512, 2008.

[7] K. Ball, D. B. Berch, K. F. Helmers, J. B. Jobe, M. D. Leveck, M. Marsiske, J. N. Morris, G. W. Rebok, D. M. Smith, S. L. Tennstedt, *et al.*, “Effects of cognitive training interventions with older adults: a randomized controlled trial,” *Jama*, vol. 288, no. 18, pp. 2271–2281, 2002.

[8] G. Winocur, F. I. Craik, B. Levine, I. H. Robertson, M. A. Binns, M. Alexander, S. Black, D. Dawson, H. Palmer, T. Mchugh, *et al.*, “Cognitive rehabilitation in the elderly: Overview and future directions,” *Journal of the International Neuropsychological Society: JINS*, vol. 13, no. 1, p. 166, 2007.

[9] E. M. Zelinski, “Far transfer in cognitive training of older adults,” *Restorative neurology and neuroscience*, vol. 27, no. 5, pp. 455–471, 2009.

[10] S. L. Willis, S. L. Tennstedt, M. Marsiske, K. Ball, J. Elias, K. M. Koepke, J. N. Morris, G. W. Rebok, F. W. Unverzagt, A. M. Stoddard, *et al.*, “Long-term effects of cognitive training on everyday functional outcomes in older adults,” *Jama*, vol. 296, no. 23, pp. 2805–2814, 2006.

[11] S. M. Barnett and S. J. Ceci, “When and where do we apply what we learn?: A taxonomy for far transfer.,” *Psychological bulletin*, vol. 128, no. 4, p. 612, 2002.

[12] K. M. Haut, K. O. Lim, and A. MacDonald, “Prefrontal cortical changes following cognitive training in patients with chronic schizophrenia: effects of practice, generalization, and specificity,” *Neuropsychopharmacology*, vol. 35, no. 9, pp. 1850–1859, 2010.

[13] K. C. Kirasic, “Spatial cognition and behavior in young and elderly adults: implications for learning new environments.,” *Psychology and aging*, vol. 6, no. 1, p. 10, 1991.

[14] M. Lövdén, S. Schaefer, H. Noack, N. C. Bodammer, S. Kühn, H.-J. Heinze, E. Düzel, L. Bäckman, and U. Lindenberger, “Spatial navigation training protects the hippocampus against age-related changes during early and late adulthood,” *Neurobiology of aging*, vol. 33, no. 3, pp. 620–e9, 2012.

[15] E. C. Tolman, “Cognitive maps in rats and men.,” *Psychological review*, vol. 55, no. 4, p. 189, 1948.

[16] R. M. Kitchin, “Cognitive maps: What are they and why study them?,” *Journal of environmental psychology*, vol. 14, no. 1, pp. 1–19, 1994.

[17] K. Konishi and V. D. Bohbot, “Spatial navigational strategies correlate with gray matter in the hippocampus of healthy older adults tested in a virtual maze,” *Frontiers in aging neuroscience*, vol. 5, p. 1, 2013.

[18] G. L. West, B. R. Zendel, K. Konishi, J. Benady-Chorney, V. D. Bohbot, I. Peretz, and S. Belleville, “Playing super mario 64 increases hippocampal grey matter in older adults,” *PLoS One*, vol. 12, no. 12, p. e0187779, 2017.

[19] H. Fuchs and G. Bishop, “Research directions in virtual environments,” 1992.

[20] M. L. Heilig, “Sensorama simulator,” Aug. 28 1962. US Patent 3,050,870.

[21] P. Cipresso, I. A. C. Giglioli, M. A. Raya, and G. Riva, “The past, present, and future of virtual and augmented reality research: a network and cluster analysis of the literature,” *Frontiers in psychology*, vol. 9, p. 2086, 2018.

[22] M. Slater, B. Lotto, M. M. Arnold, and M. V. Sánchez-Vives, “How we experience immersive virtual environments: the concept of presence and its measurement,” *Anuario de Psicología, 2009, vol. 40, p. 193-210*, 2009.

[23] R. I. García-Betances, M. T. Arredondo Waldmeyer, G. Fico, and M. F. Cabrera-Umpiérrez, “A succinct overview of virtual reality technology use in alzheimer’s disease,” *Frontiers in aging neuroscience*, vol. 7, p. 80, 2015.

[24] S. M. Slobounov, W. Ray, B. Johnson, E. Slobounov, and K. M. Newell, “Modulation of cortical activity in 2d versus 3d virtual reality environments: an eeg study,” *International Journal of Psychophysiology*, vol. 95, no. 3, pp. 254–260, 2015.

[25] J. J. LaViola Jr, “A discussion of cybersickness in virtual environments,” *ACM Sigchi Bulletin*, vol. 32, no. 1, pp. 47–56, 2000.

[26] J. G. Casali, “Vehicular simulation-induced sickness. volume 1. an overview.,” tech. rep., VIRGINIA POLYTECHNIC INST AND STATE UNIV BLACKSBURG DEPT OF INDUSTRIAL . . . , 1985.

[27] S. Weech, S. Kenny, and M. Barnett-Cowan, “Presence and cybersickness in virtual reality are negatively related: a review,” *Frontiers in psychology*, vol. 10, p. 158, 2019.

[28] P. Belchior, “Cognitive training with video games to improve driving skills and driving safety among older adults,” *University of Florida*, 2007.

[29] V. Manera, P.-D. Petit, A. Derreumaux, I. Orvieto, M. Romagnoli, G. Lyttle, R. David, and P. H. Robert, “‘kitchen and cooking,’a serious game for mild cognitive impairment and alzheimer’s disease: a pilot study,” *Frontiers in aging neuroscience*, vol. 7, p. 24, 2015.

[30] H. B. Abdessalem, M. Boukadida, M.-A. Bruneau, P. Robert, S. Belleville, R. David, and C. Frasson, “Immersion en train thérapeutique pour la relaxation de patients alzheimer,” *French Journal of Psychiatry*, vol. 1, p. S152, 2019.

[31] S. H. Fairclough, “Physiological computing and intelligent adaptation,” in *Emotions and Affect in Human Factors and Human-Computer Interaction*, pp. 539–556, Elsevier, 2017.

[32] A. A. Rizzo, J. G. Buckwalter, T. Bowerly, C. Van Der Zaag, L. Humphrey, U. Neumann, C. Chua, C. Kyriakakis, A. Van Rooyen, and D. Sisemore, “The virtual classroom: a virtual reality environment for the assessment and rehabilitation of attention deficits,” *CyberPsychology & Behavior*, vol. 3, no. 3, pp. 483–499, 2000.

[33] G. Optale, C. Urgesi, V. Busato, S. Marin, L. Piron, K. Priftis, L. Gamberini, S. Capodieci, and A. Bordin, "Controlling memory impairment in elderly adults using virtual reality memory training: a randomized controlled pilot study," *Neurorehabilitation and neural repair*, vol. 24, no. 4, pp. 348–357, 2010.

[34] P. Gamito, J. Oliveira, C. Coelho, D. Morais, P. Lopes, J. Pacheco, R. Brito, F. Soares, N. Santos, and A. F. Barata, "Cognitive training on stroke patients via virtual reality-based serious games," *Disability and rehabilitation*, vol. 39, no. 4, pp. 385–388, 2017.

[35] F. Imbeault, B. Bouchard, and A. Bouzouane, "Serious games in cognitive training for alzheimer's patients," in *2011 IEEE 1st International Conference on Serious Games and Applications for Health (SeGAH)*, pp. 1–8, IEEE, 2011.

[36] G. Caggianese, A. Chirico, G. De Pietro, L. Gallo, A. Giordano, M. Predazzi, and P. Neroni, "Towards a virtual reality cognitive training system for mild cognitive impairment and alzheimer's disease patients," in *2018 32nd International Conference on Advanced Information Networking and Applications Workshops (WAINA)*, pp. 663–667, IEEE, 2018.

[37] S. M. Fisch, "Making educational computer games" educational"," in *Proceedings of the 2005 conference on Interaction design and children*, pp. 56–61, 2005.

[38] V. Terzis, C. N. Moridis, and A. A. Economides, "Measuring instant emotions based on facial expressions during computer-based assessment," *Personal and ubiquitous computing*, vol. 17, no. 1, pp. 43–52, 2013.

[39] J. De Lemos, G. R. Sadeghnia, Í. Ólafsdóttir, and O. Jensen, "Measuring emotions using eye tracking," in *Proceedings of measuring behavior*, vol. 226, pp. 225–226, 2008.

[40] K. M. Gilleade and A. Dix, "Using frustration in the design of adaptive videogames," in *Proceedings of the 2004 ACM SIGCHI International Conference on Advances in computer entertainment technology*, pp. 228–232, 2004.

[41] R. Ghali, H. B. Abdessalem, and C. Frasson, "Improving intuitive reasoning through assistance strategies in a virtual reality game," in *The Thirtieth International Flairs Conference*, 2017.

[42] L. Clare and R. T. Woods, "Cognitive training and cognitive rehabilitation for people with early-stage alzheimer's disease: A review," *Neuropsychological rehabilitation*, vol. 14, no. 4, pp. 385–401, 2004.

[43] A. Morley, L. Hill, and A. Kaditis, "10–20 system eeg placement," *Eur. Respir. Soc*, 2016.

[44] R. Maskeliunas, R. Damasevicius, I. Martisius, and M. Vasiljevas, "Consumer-grade eeg devices: are they usable for control tasks?," *PeerJ*, vol. 4, p. e1746, 2016.

[45] T. Harrison, "The emotiv mind: Investigating the accuracy of the emotiv epoc in identifying emotions and its use in an intelligent tutoring system," 2013.

[46] T. D. Pham and D. Tran, "Emotion recognition using the emotiv epoc device," in *International Conference on Neural Information Processing*, pp. 394–399, Springer, 2012.

[47] T. Matlovic, P. Gaspar, R. Moro, J. Simko, and M. Bielikova, "Emotions detection using facial expressions recognition and eeg," in *2016 11th International Workshop on Semantic and Social Media Adaptation and Personalization (SMAP)*, pp. 18–23, 2016.

[48] Z. S. Nasreddine, N. A. Phillips, V. Bédirian, S. Charbonneau, V. Whitehead, I. Collin, J. L. Cummings, and H. Chertkow, "The montreal cognitive assessment, moca: a brief screening tool for mild cognitive impairment," *Journal of the American Geriatrics Society*, vol. 53, no. 4, pp. 695–699, 2005.

[49] D. Watson, L. A. Clark, and G. Carey, "Positive and negative affectivity and their relation to anxiety and depressive disorders.," *Journal of abnormal psychology*, vol. 97, no. 3, p. 346, 1988.

[50] R. S. Kennedy, N. E. Lane, K. S. Berbaum, and M. G. Lilienthal, “Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness,” *The international journal of aviation psychology*, vol. 3, no. 3, pp. 203–220, 1993.

[51] C. Lallemand, V. Koenig, G. Gronier, and R. Martin, “Création et validation d’une version française du questionnaire attrakdiff pour l’évaluation de l’expérience utilisateur des systèmes interactifs,” *European Review of Applied Psychology*, vol. 65, no. 5, pp. 239–252, 2015.

[52] C. J. Watkins and P. Dayan, “\cal q-learning,” *Machine learning*, vol. 8, no. 3-4, pp. 279–292, 1992.

[53] V. Mnih, K. Kavukcuoglu, D. Silver, A. A. Rusu, J. Veness, M. G. Bellemare, A. Graves, M. Riedmiller, A. K. Fidjeland, G. Ostrovski, *et al.*, “Human-level control through deep reinforcement learning,” *nature*, vol. 518, no. 7540, pp. 529–533, 2015.

Appendix A

Accepted Paper in 16th International Conference on Intelligent Tutoring Systems, 2020

Conference

Intelligent Tutoring Systems ,16th International Conference, ITS 2020, Athens, Greece, June 8–12, 2020

Part of the Lecture Notes in Computer Science book series (LNCS, volume 12149). Also part of the Programming and Software Engineering book sub series (LNPSE, volume 12149).

Author Contribution

Researchers at the University of Montreal

Alexie Byrns: Original abstract idea of using cognitive maps in orientation based games as cognitive exercise based on the article by Greg et al. [**18**].
Claude Frasson: Development of idea to a feasible product. Supporting and funding and reviewing the project.
Marwa Boukadida: Main developer of the VR environment. Performing the experiments with participants at IUGM.
Hamdi Ben Abdessalem: Support developer of VR environment. Integration and testing of the developed product. Development of memory and attention exercises. Performing the experiments with participants at IUGM. Analysis of results.
Manish Kumar Jha: Support developer of VR environment. Developed idea of using hints to help participants to solve the tasks. Developed rule based guidance system

and improvement using reinforcement learning. Integration and testing of the developed product. Analysis of results. Main author of the article 'Improving Cognitive and Emotional StateUsing 3D Virtual Reality Orientation Game' accepted at ITS 2020.

Researchers at Institut universitaire de gériatrie de Montréal

Marc Cuesta: Performing the experiments with participants at IUGM.
Marie-Andrée Bruneau: Performing the experiments with participants at IUGM.
Sylvie Belleville : Reviewing the work. Performing the experiments with participants at IUGM.

Editors

Vivekanandan Kumar , Full Professor at the Athabasca University.
Christos Troussas, Adjunct Lecturer at the University of West Attica (Greece)

Improving Cognitive and Emotional State Using 3D Virtual Reality Orientation Game

Manish Kumar Jha[1], Marwa Boukadida[1],
Hamdi Ben Abdessalem[1(✉)], Alexie Byrns[1], Marc Cuesta[2],
Marie-Andrée Bruneau[2], Sylvie Belleville[2], and Claude Frasson[1]

[1] Département d'Informatique et de Recherche Opérationnelle, Université de Montréal, Montréal, Canada
{manish.jha,marwa.boukadida,hamdi.ben.abdessalem,alexie.byrns}@umontreal.ca, frasson@iro.umontreal.ca

[2] Centre de Recherche de l'Institut de Gériatrie de Montréal, Montréal, Canada
marc.cuesta@criugm.qc.ca, {marie.andree.bruneau,sylvie.belleville}@umontreal.ca

Abstract. Patients suffering from Alzheimer's Disease (AD) exhibit an impairment in performing tasks related to spatial navigation. Tasks which require navigational skills by building a cognitive map of the surrounding are found effective in cognitive training. In this paper we investigated the effect of cognitive training using a fully immersive 3D VR orientation game. We implemented an intelligent guidance system which helps to reduce the negative emotions if the participants experience difficulty completing the quests of the game. We found that after playing the orientation game, participants performed better in memory and in certain attention exercises. We also studied the effects of guidance system to reduce the frustration during cognitive training using VR environments.

Keywords: Virtual Reality · Orientation · EEG · Immersive environment · Game adaptation · Assistance system

1 Introduction

The ability to find one's way using spatial reference frames and cues from our surrounding is a complex cognitive process. Studies show that navigational skills decline with ageing [1]. However, people with Alzheimer's disease (AD) display a significantly higher decline, which is also one of the early symptoms of the disease. Research shows that spatial navigation training programs in older persons led to improvements in spatial performances [2].

Virtual Reality (VR) applications can be used to address the challenges of cognitive training of dementia patients due to a high level of interaction possible within a virtual environment (VE) without being in any risk otherwise posed by real-life surroundings. In this paper, we present a fully immersive VE where the participant must find items of interest in a public garden. However, navigating in an unfamiliar place can be challenging, consequently leading to higher negative emotions and a tendency to give up

V. Kumar and C. Troussas (Eds.): ITS 2020, LNCS 12149, pp. 220–225, 2020.
https://doi.org/10.1007/978-3-030-49663-0_26

before completing the experiment. Research shows that it is more beneficial for patients with cognitive impairment to be helped through the completion of a challenge, rather than see the challenge failed [3]. It is important to present both audio and visual cues to cater to the needs of a specific profile of patients suffering from either visual or auditory impairments [4]. Thus, real-time assistance with audio and visual feedback is one of the mandatory components in games for elders, to incorporate mechanism which achieves high-level engagement by keeping player filled with positive emotions. Hence, we implemented an intelligent guidance system based on participant's behavior, that helps them to complete the tasks without being explicitly asked for help.

Our goal for this study is to investigate the effect of the orientation game and the guidance system on the cognitive functions of people suffering from subjective cognitive decline (SCD), a preclinical state of possible Alzheimer's Disease (AD). We state our research objectives as the following: **Q1: is it possible to stimulate the brain using a virtual maze game in order to enhance attention and memory? Q2: is it possible to help participants in order to reduce negative emotions?**

The rest of this paper is organized as follows. In Sect. 2, we discuss the related works and the cognitive map theory. In Sect. 3, we describe the orientation game: the environment, the objectives and the guidance system. In Sect. 4, we detail the cognitive tests and the experimental procedure undertaken to validate our hypotheses. Finally, in Sect. 5 we present and discuss the obtained results.

2 Related Works

According to the cognitive map theory, the formation of representations of spatial information – in other words, the creation of a cognitive map – helps reduce cognitive load and increases recall and the encoding of novel information [5, 6]. Studies report that decreases in the volume of the hippocampus – a structure playing a key role in memory – correlate with a decline in cognitive function. Indeed, it is speculated that increasing grey matter in the hippocampus could entail better memory [7]. Interestingly, playing 3D video games over a period, such as Super Mario 64, reportedly increases hippocampal volume [8], as well as increases performance in episodic and spatial memory quests. It is speculated that 3D games, such as Super Mario 64, lead players to create a cognitive map of the environment.

It has been observed that virtual reality environments (VE) lead to formation of cognitive maps similar to the real environments [9]. This makes VR games ideal for simulating real life scenarios for cognitive training of elderly. Certain VR applications have concentrated on games focused on performing activities of daily life such as cooking, driving and shopping [10]. Another key advantage of using VE is that they offer a safe way to achieve high level of interaction adaptable to the characteristics and needs of individual patients [10, 11]. A fully immersive VE offers higher sense of 'presence' and interaction which subsequently affects the behavioral responses of patients.

3 Orientation Game

3.1 Environment: Orientation Game

The fully immersive VR environment simulates a botanical garden in the form of a 5×5 maze. In this environment, trees form the walls of the maze and clearings through the trees are the pathways. The participant starts at one end of the garden and has to navigate using a joystick by clicking in the direction in which he/she intends to move. Other elements in the game are: 1) a map of the garden with geographical directions, 2) the position and direction of the user shown by a red arrow in the map, 3) a flashing blue circle representing the location of the items, 4) visual hints displayed when needed and, 5) verbal messages to the participant.

The game starts with a tutorial to let the users familiarize themselves with the environment and the controls. It consists of four quests. For the first three quests, the participant is asked to collect an item located at a specific location of the 5×5 maze. We display the name of the item and its location by a flashing blue circle on the map for 5 s. The user needs to reach the location and collect the requested item. When the item is collected, we remove it from the list and display the next item and its location.

3.2 Guidance System

We implemented a rule-based guidance system that provides navigational hints or audio and visual messages to the participants and helps them in completing the quests. It actively monitors the participant's emotions, namely frustration, excitement, engagement, meditation, and valence using Emotiv electroencephalograph (EEG) headset data in real-time. On sensing a situation where the participant may need a hint, it sends a message to the VR system, which displays the hints in form of location in the map or audio and visual messages in case of text-based hints. Figure 1 shows the different hints provided by the guidance system.

Hint levels		Participant's position	Object's location	Message (Audio and Visual)	Highlighted Path in Map
Level 1		✓	✓		
Level 2	2-1	✓	✓	*Please check the position of object on the map.*	
	2-2	✓	✓	*You are too far. Try to take few steps back.*	
	2-3	✓	✓	*1. Good job! Almost there.* *2. Keep up the good work* *3. You are going in right direction*	
Level 3		✓		*Follow the displayed path. The object is somewhere nearby.*	Path to cell nearest to target object
Level 4		✓		*Follow the given path to find the object*	Complete path to the object.

Fig. 1. Different level of hints

The hints provided by the guidance system are activated in three different cases:

1. Emotions: At every second, the mean of the change and the rate of the change of emotion values in past ten seconds are used to calculate a net score for each

emotion. The emotion with the maximum score is compared with an empirically defined threshold to activate the emotion-based hints.
2. Away from target: If the participant takes three steps or more, all of which are at four blocks or more from the target, the map displaying target location is activated.
3. No Movement: If the participant doesn't move for more than a given amount of time, the map displaying target location is activated.

The details change with different levels of hints provided by the guidance system. Level 1 provides the least information and displays only the participant's location and the object's location on the map. Additionally, level 2 displays a text message in a prompt in the VR environment along with the verbal narration of the message. Level 3 hint highlights a path in the map, which the participant can follow to reach a location immediately next to the actual location of the object. This leaves some scope of exploration and the participant needs to search for the object in all possible directions. Level 4 hint highlights the complete path leading to the object's location on the map.

In case of activation based on participant's emotion, if the hint is triggered by frustration, the level of hint increases which provides more details to find the object. On the other hand, if the hint is triggered by positive emotions such as excitement or engagement, the level of hint decreases. In case of hint activated due to no movement of the participant, the level increases every fifteen seconds till the participant moves. When the participant is away from the object as determined by the guidance system, the hint provided is always level 2-1.

4 Experiments

Since our main goal is to analyze the impact of orientation therapy and building a cognitive map on the attention and memory performances, we developed 3 attention exercises and 3 memory exercises to compare participants' performances before and after playing the game. We tested our approach with 15 participants (11 females) with subjective cognitive decline (SCD) and a mean age = 73.4 (SD = 5.73).

The participants took part in two sessions. In the first session, we performed some assessments to ensure that they were eligible to conduct the experiments. During the second session (experimental session), the participants were invited to fill a pre-experiment form. Afterwards, we equipped them with an EEG headset, and they start resolving attention and memory exercises. When they complete the exercises, we equip them with Fove VR headset, and they start the Orientation Game. Following the game, we removed the Fove VR headset and the participants were asked to complete the attention and memory exercises again but with different examples. Finally, we remove the EEG headset and the fill up a post-experiment form.

5 Results and Discussion

In order to study the enhancement in attention and memory, we analyzed performance improvement before and after the orientation therapy of the first three exercises (Attention exercises). On exercise 1, the general mean improvement was 6.67%. on the second exercise, there was a mean improvement of 0.61%. And the performance improvement of the third exercise was 0%. We also analyzed the performance improvement before and after the orientation therapy of the memory exercises (exercise 4, 5 and 6). For the fourth exercise, the mean improvement was 1.11%. For the fifth exercise, the mean improvement was 12%. Finally, the mean improvement is 26.67% for exercise 6 which is the highest percentage of improvement. These results show increase of memory performance following the orientation therapy and an improvement in attention abilities in certain participants.

Next, we analyzed the frustration of the participants before and after the hints are provided. As shown in Fig. 2, the guidance system provided at least one or more hints to 9 out of the 15 participants. In the rest of the cases, participants didn't have any difficulty in completing the quests.

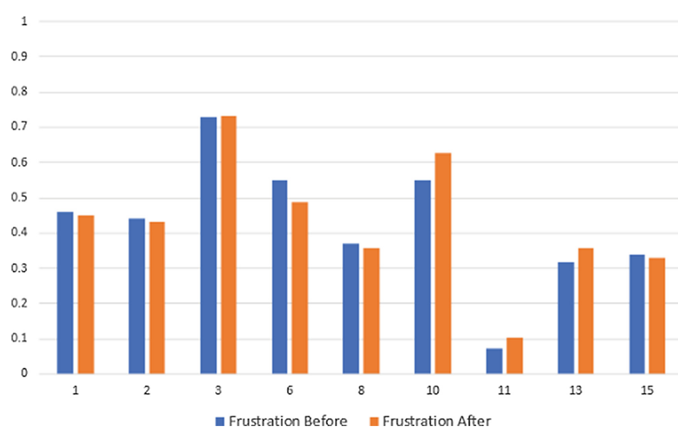


Fig. 2. Average values of frustrations of each participant, ten seconds before and after the hints

We notice that in 6 out of 9 participants, frustration values in next ten seconds decreased for the hints provided to them. Participants 8, 10, 11 and 13 were respectively provided one, two, one, and one hint, while the rest of the participants received at least five or more hints. For the different types of hints, we observed that except for hint level 2-2 and hint level 4, the average values of frustration for all the participants were lesser in the next ten seconds after the hints were provided. Hint level 2-2 provides a warning message: 'You're too far. Try to take few steps back.'. This led the participants to believe that they might be doing something wrong leading to higher frustration. Hint level 4 displays the complete path to the item's location. But, since we configured the hint to appear for only four seconds, it was not enough for the participants to memorize the complete path, which may lead to a more frustration.

6 Conclusion

In this paper, we designed a 3D VR orientation game in with real-time guidance system which can be used for cognitive training of the patients suffering from pre-clinical states of Alzheimer's disease. The results show an improvement in memory performance for most of the participants, and better attention abilities for some of the participants after the therapy. The guidance system is effective in reducing the frustration while solving quests of the game. In some cases, the decrease is not significant, but the stabilization of frustration after a continuous increase after the hint is provided shows the usefulness of hints. Also, the increase in negative emotions after two hints shows that hints need to be carefully designed to give positive messages and the time taken to understand the hints should be taken in consideration.

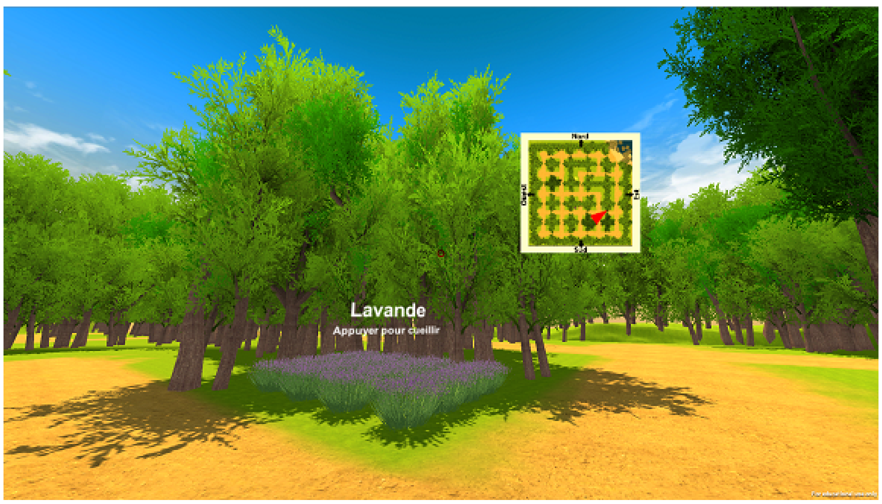
Acknowledgement. We acknowledge NSERC-CRD and Beam Me Up for funding this work.

References

1. Kirasic, K.C.: Spatial cognition and behavior in young and elderly adults: implications for learning new environments. Psychol. Aging **6**, 10 (1991)
2. Lövdén, M., et al.: Spatial navigation training protects the hippocampus against age-related changes during early and late adulthood. Neurobiol. Aging **33**, 620-e9 (2012)
3. Pigot, H., Mayers, A., Giroux, S.: The intelligent habitat and everyday life activity support. In: Proceedings of the 5th International conference on Simulations in Biomedicine, April 2003
4. Marin, J.G., Navarro, K.F., Lawrence, E.: Serious games to improve the physical health of the elderly: a categorization scheme. In: International Conference on Advances in Human-oriented and Personalized Mechanisms, Technologies, and Services. Barcelona, Spain (2011)
5. Kitchin, R.M.: Cognitive maps: What are they and why study them? J. Environ. Psychol. **14**, 1–19 (1994)
6. Tolman, E.C.: Cognitive maps in rats and men. Psychol. Rev. **55**, 189 (1948)
7. Konishi, K., Bohbot, V.D.: Spatial navigational strategies correlate with gray matter in the hippocampus of healthy older adults tested in a virtual maze. Front. Aging Neurosci. **5**, 1 (2013)
8. West, G.L., et al.: Playing Super Mario 64 increases hippocampal grey matter in older adults. PLoS ONE **12**, e0187779 (2017)
9. Péruch, P., Gaunet, F.: Virtual environments as a promising tool for investigating human spatial cognition. Cahiers de Psychologie Cognitive/Curr. Psychol. Cogn. (1998)
10. Ball, K., et al.: Effects of cognitive training interventions with older adults: a randomized controlled trial. JAMA **288**, 2271–2281 (2002)
11. Imbeault, F., Bouchard, B., Bouzouane, A.: Serious games in cognitive training for Alzheimer's patients. In: 2011 IEEE 1st International Conference on Serious Games and Applications for Health (SeGAH) (2011)

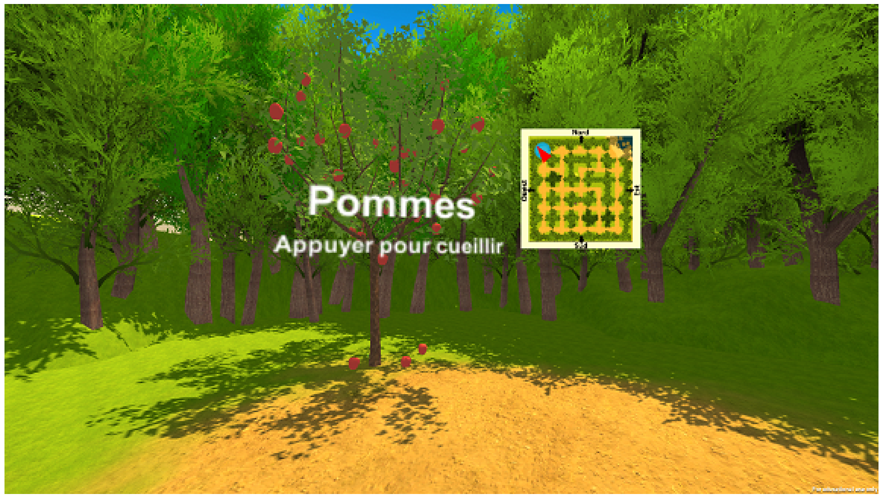
Appendix B

VR Scenes of Items in the Orientation Game

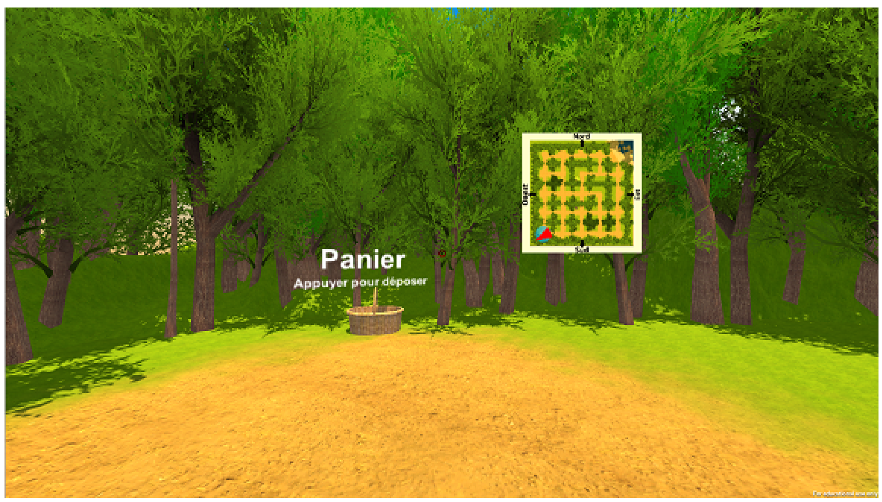


(a) Lavender

(b) Coconuts



(c) Apples

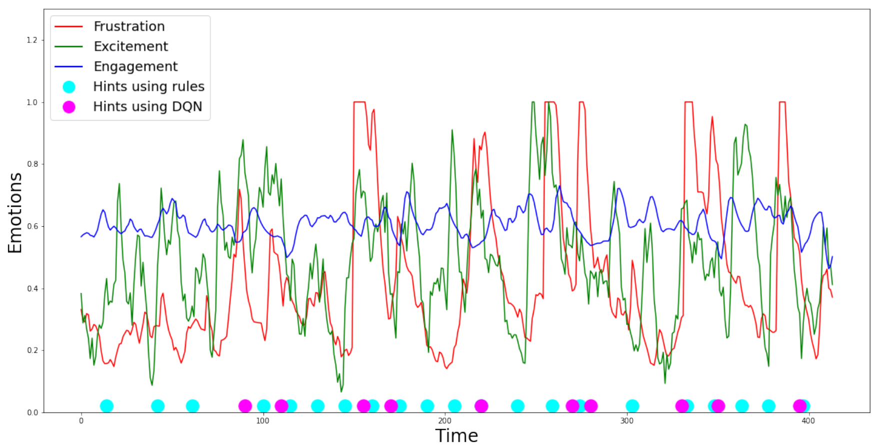


(d) Basket

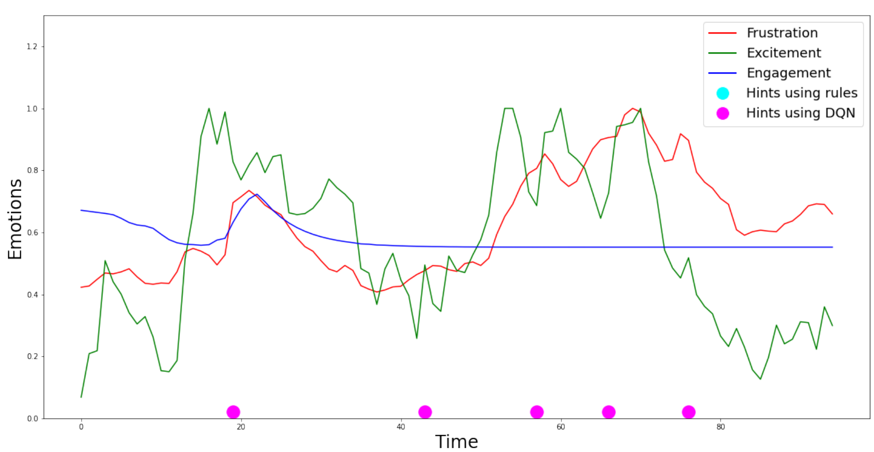
Figure B.0. VR Scenes of Items to be Collected in the Orientation Game

Appendix C

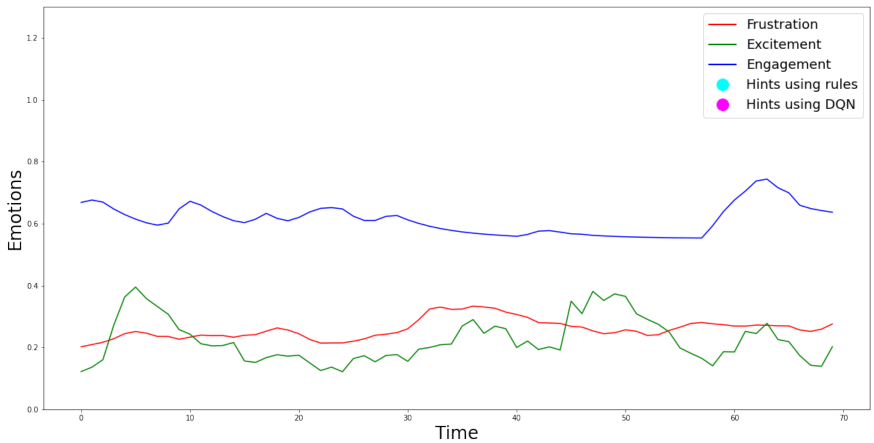
Comparison of Hints Provided by Rule-based vs RL Guidance System



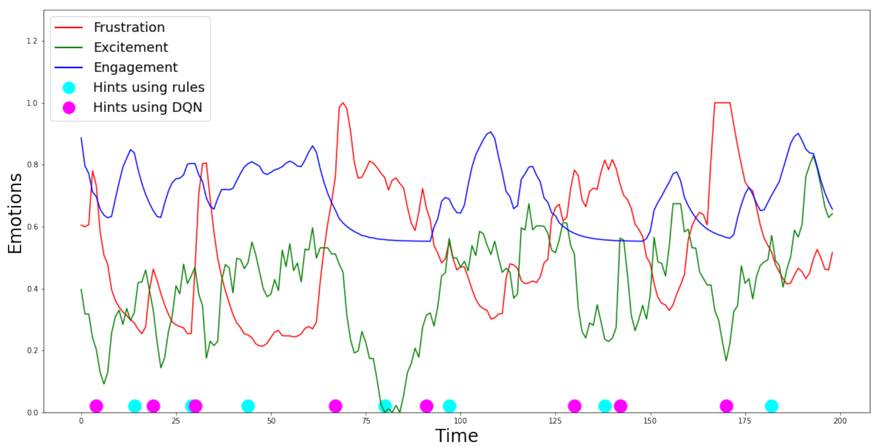
(a) P1



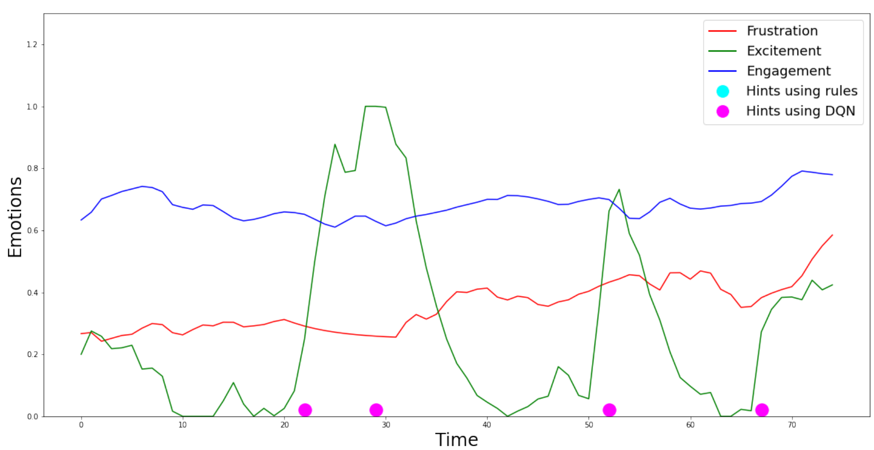
(b) P2


Frustration
Excitement
Engagement
Hints using rules
Hints using DQN
Emotions
Time

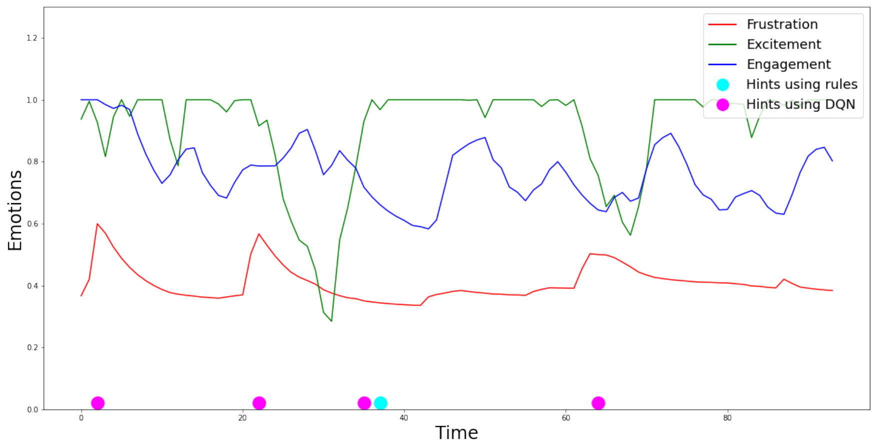
(c) P3

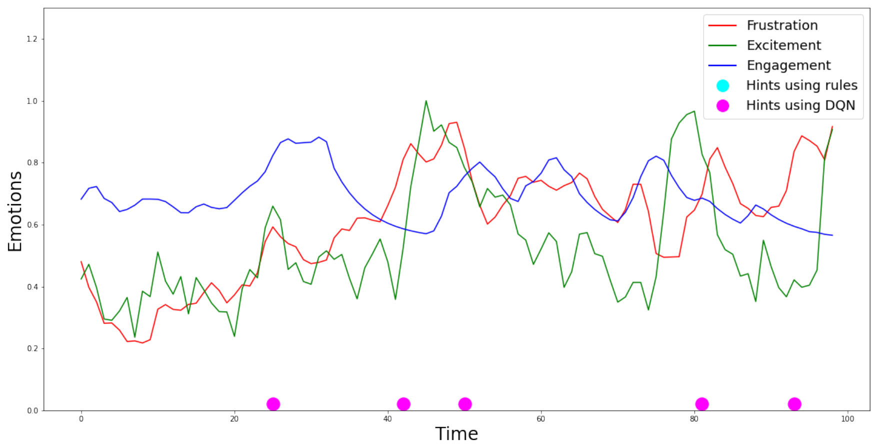

Frustration
Excitement
Engagement
Hints using rules
Hints using DQN
Emotions
Time

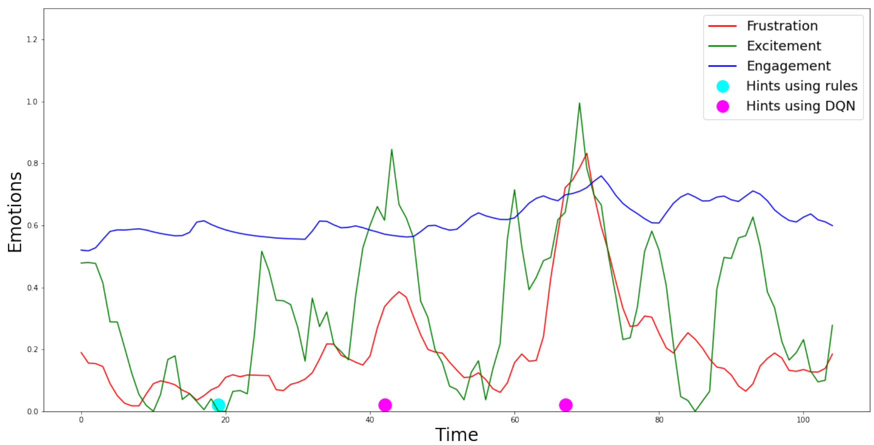
(d) P4

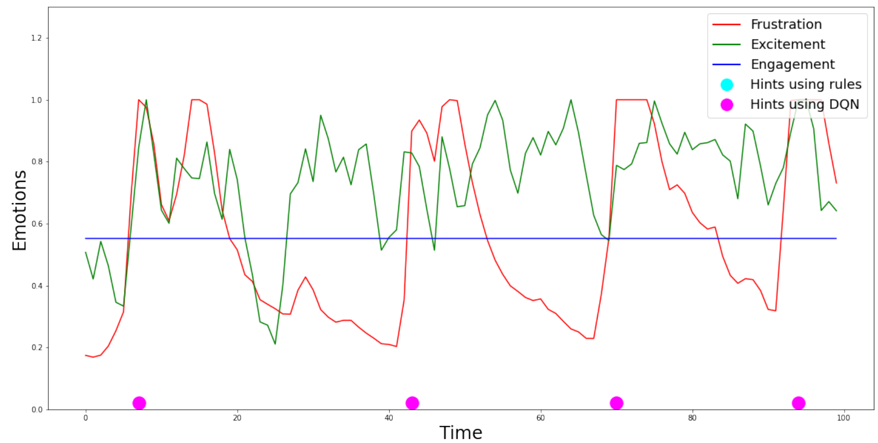

Frustration
Excitement
Engagement
Hints using rules
Hints using DQN
Emotions
Time

(e) P5

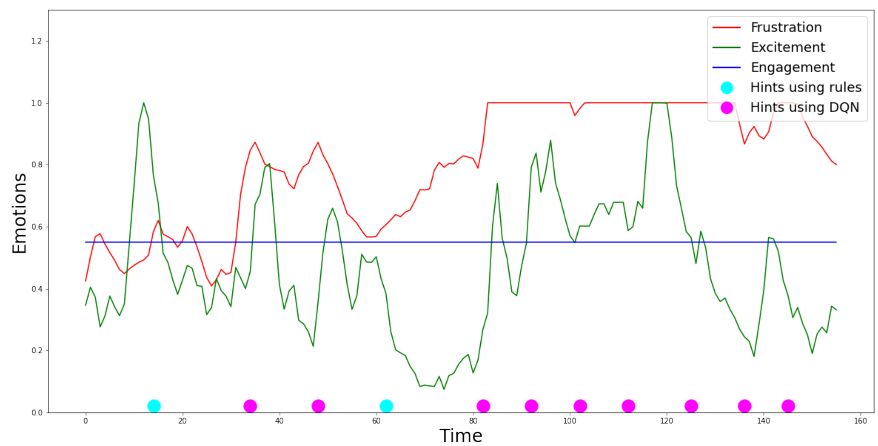

Frustration
Excitement
Engagement
Hints using rules
Hints using DQN
Emotions
Time
(f) P6


Frustration
Excitement
Engagement
Hints using rules
Hints using DQN
Emotions
Time
(g) P7

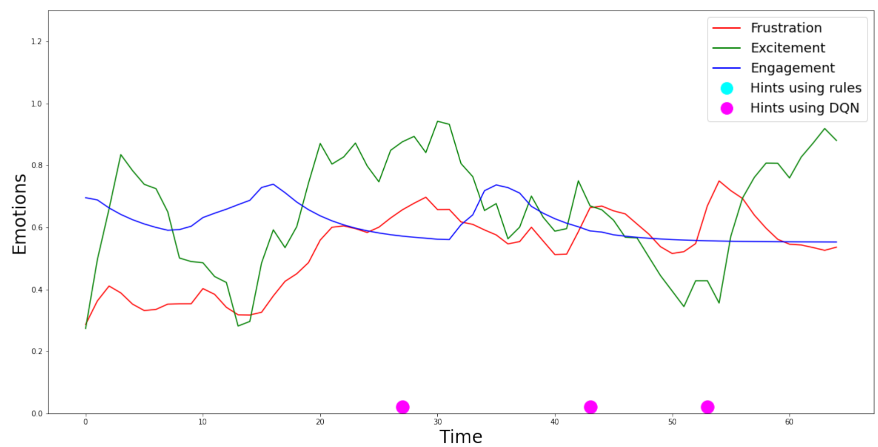

Frustration
Excitement
Engagement
Hints using rules
Hints using DQN
Emotions
Time
(h) P8



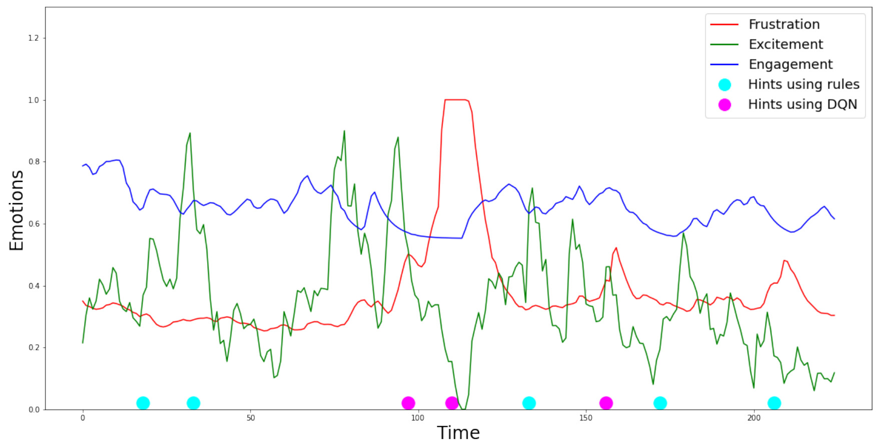
(i) P9



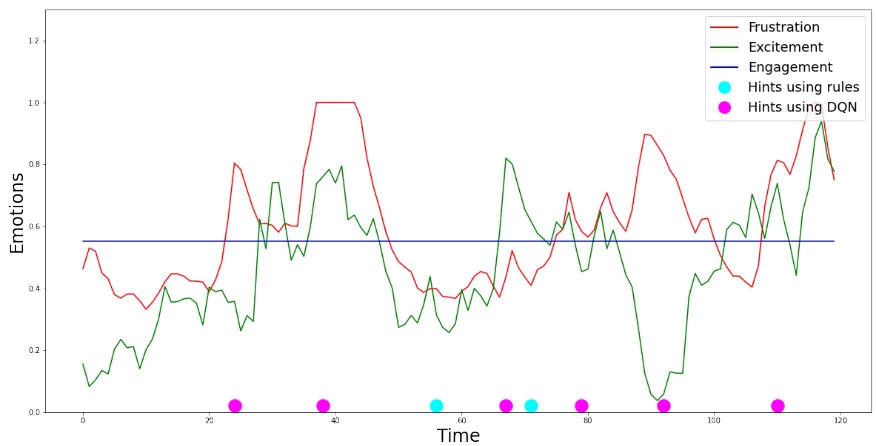
(j) P10



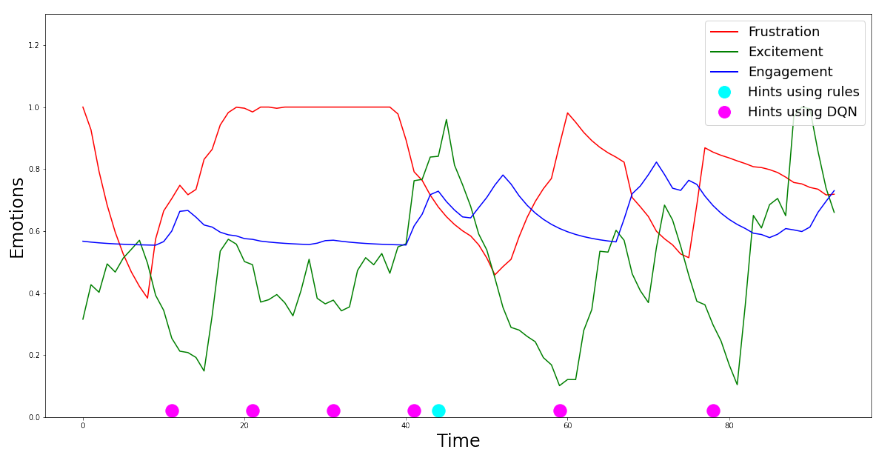
(k) P11



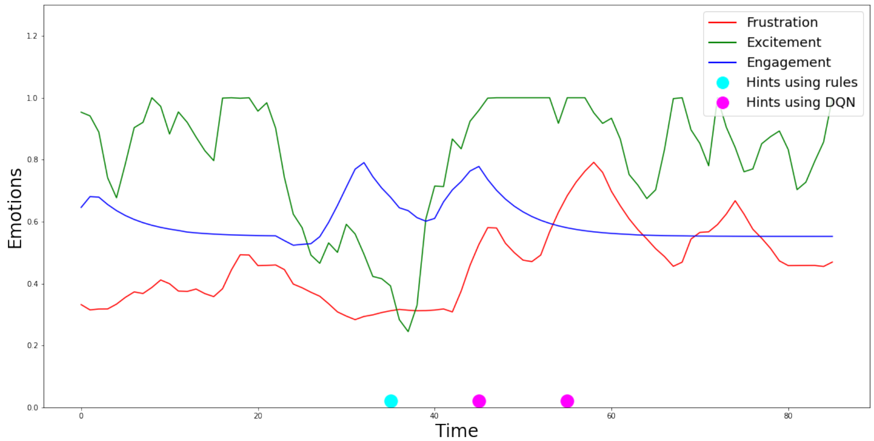
(l) P12



(m) P13



(n) P14



(o) P15

Figure C.0. Comparison of timestamp to provide hints to participants by the two systems

Appendix D

Post Experiment Questionnaire

S. No.	**Survey Questions**					
1	Est-ce que vous vous êtes immergé(e) dans l'environnement de réalité virtuelle ? How immersive do you find the virtual reality environment?	1	2	3	4	5
2	Est-ce que la réalité virtuelle a eu un impact positif (par rapport à une interaction avec un écran 2D) ? Did 3D VR environment have a positive impact on you compared to a 2D environment?	1	2	3	4	5
3	Est-ce que vous avez eu de la difficulté à vous déplacer dans l'environnement virtuel ? Did you have difficulty moving around in the virtual environment?	Oui Yes	Non No			
4	Est-ce que vous avez eu besoin d'aide (carte, messages, chemins) pour vous orienter dans l'environnement virtuel ? Did you need help (map, messages, paths) to orient yourself in the virtual environment?)	Oui Yes	Non No			
5	Est-ce que l'aide (carte, messages, chemins) vous a été utile ? (Was the help (map, messages, paths) useful to you?)	Oui Yes	Non No			

S. No.	Survey Questions					
6	Quel était votre niveau de stress/frustration/confusion avant de recevoir l'aide (carte, messages, chemins)? What was your level of stress/frustration/confusion before receiving help (map, messages, paths)?	1	2	3	4	5
7	Quel était votre niveau de stress/frustration/confusion après avoir reçu l'aide (carte, messages, chemins...)? What was your level of stress/frustration/confusion after receiving help (map, messages, paths ...)?	1	2	3	4	5
8	Est-ce que vous étiez stressé(e) durant les exercices? (avant le jeu d'orientation) Were you stressed during the exercises before the orientation game?	Oui Yes	Non No			
9	Est-ce que vous étiez stressé(e) durant les exercices (après le jeu d'orientation) ? Were you stressed during the exercises after the orientation game?	Oui Yes	Non No			
10	Si le cas se présente, désireriez vous participer à d'autres expériences similaires ? If applicable, would you like to participate in other similar experiences?	Oui Yes	Non No			