UNIVERSITY OF CAPE TOWN
DEPARTMENT OF COMPUTER SCIENCE

CS/IT Honours
Final Paper 2020

Title: The Effect of Interaction on Eliciting Sadness in Virtual Reality

Author: Brent Lee van der Walt

Project Abbreviation: VREmote

Supervisor(s): Dr. James Gain, Dr. Gosia Lipinska, Siphumelele Sigwebela

Category	*Min*	*Max*	Chosen
Requirement Analysis and Design	*0*	*20*	15
Theoretical Analysis	*0*	*25*	0
Experiment Design and Execution	*0*	*20*	5
System Development and Implementation	*0*	*20*	20
Results, Findings and Conclusions	*10*	*20*	10
Aim Formulation and Background Work	*10*	*15*	10
Quality of Paper Writing and Presentation	*10*		10
Quality of Deliverables	*10*		10
Overall General Project Evaluation (*this section allowed only with motivation letter from supervisor*)	*0*	*10*	0
Total marks		**80**	

The Effect of Interaction on Eliciting Sadness in Virtual Reality

Brent van der Walt
Computer Science Department
University of Cape Town
vwlbre001@myuct.ac.za

ABSTRACT

Interaction, the degree to which users of a medium can influence the content of the environment, is one of the key features of Virtual Reality, yet there has been surprisingly little research on its effect in evoking emotions. Virtual Reality technologies are advancing at a rapid pace and becoming more readily available for wider use, so finding ways to improve this experience, either for therapy or entertainment is important. This paper focuses on the creation of an environment that elicits sadness, with the ability to manipulate the amount of interaction between the participant and the environment, as well as the use of Artificial Intelligence that reacts to the participant's actions by dynamically altering its behavior. This environment and artificial intelligence was evaluated by experts in the field of Virtual Reality through heuristic evaluation. The environment was able to manipulate the amount of interaction that occurred successfully with few flaws in how it was executed, and the artificial intelligence was believable and dynamic but was not successful in having enough variation in the behaviors. Our results indicate that the environment is suitable for use in future iterations, but that the artificial intelligence would need to be reworked if variation in behaviors should be the focus.

KEYWORDS

Virtual Reality, Virtual Environments, Interaction, Mental Model, Parasympathetic Activation, Artificial Intelligence

1 Introduction

Mental disorders such as mood and anxiety disorders are problems that are, and will continue to be, prevalent in the current world. One of these such issues, with 322 million cases [1], is depression. The virtual world can confer many therapeutic benefits over naturalistic or traditional therapeutic situations, because it can provide a controlled environment where feelings of sadness can be elicited for purposes such as understanding how an individual experiences and reacts in sad situations. The reactions of people with disorders can be compared to those without or to previous reactions, providing knowledge on the nature of the emotion and allowing testing as to whether certain interventions are working. Virtual Reality (VR), the use of technology to create a simulated environment, has been shown to be an effective medium for psychotherapeutic techniques [2, 3, 4] and has been used to treat grief [5]. VR can evoke the same reactions and emotions as a real experience in a stressor environment [6] with the ability to manipulate the environment of the patient in an easy and safe way. Virtual Reality has also started to become more prevalent in the entertainment industry and the elicitation of emotion can help in storytelling, making VR applicable in this context as well.

One of the advantages of using VR for rehabilitation compared to purely visual mediums, such as video, is the ability for interaction, the degree to which users of a medium can influence the virtual environment (VE), yet there has surprisingly been minimal research on its effect in evoking emotions. Being able to manipulate emotions not only gives us a better understanding of the psychology around them, but also shows us how we can provide better therapeutic treatment or better entertainment experiences.

Apart from interaction, when agents are present in a VE it is important to have a believable Artificial Intelligence (AI), which in this context is the use of algorithms to generate responsive and adaptive behaviors in non-player characters. For the behaviors to be believable, they need to be emergent, where the AI has different reactions to events depending on the context of the environment and the user's behavior. These emergent behaviors are vital for making the reactions of the VE seem more realistic. This brings us to the aims of this project:
First, to create a VE which can manipulate the amount of interaction between the participant and the environment. Second, to create an AI in the VE that reacts to the participant's actions by dynamically altering the environment or the AI's behavior.

This paper will focus on the elicitation of sadness, due to the already existing environment which was created in a previous iteration of this project and because it is easier to measure the elicitation of a single emotion. Sadness was chosen due to its high prevalence and the possible therapeutic benefits that can be achieved by understanding it better. This report consists of 6 sections: Section 2 is an overview of background and related work; Section 3 highlights and describes the previous VE and the implementation of the interaction and AI in the new VE; Section 4 describes the testing design and procedures; Section 5 is a discussion and summary of the results from the testing and describes the limitations of the project; and Section 6 details the conclusions and possible extensions for future work.

2 Background and Related Work

This section is broken into three sections, the first focusing on presence in VR, the mental model and how interaction helps form this; The second focusing on the subtype of sadness we hope to elicit and how to elicit it; The third focusing on the types of AI that are involved in this project.

2.1 Presence and Interaction

Presence and immersion have had various definitions over a large variety of papers [6, 7, 8, 9, 10], but this paper will look at them as interchangeable making use of presence. This forgoes immersion, as an objective description of the technology used to create this experience.

Presence is similar to flow - being completely absorbed and immersed in an activity such that one forgets their current physical surroundings and time [9]. The importance of presence in VR in this field is that there has been evidence to suggest that it helps evoke emotion [6, 7, 11], a directly proportional relationship has been found between presence and emotional response with a moderate effect size [6]. This means that higher presence results in a stronger elicitation of emotion and a lack of presence results in a low elicitation of emotion, making it is a necessary factor for this project.

In VR, interaction involves both actions from the participant as well as reactions from the virtual environment. These actions and reactions are the actual interactions that take place between a user and the world and which make up a part of the potential interactions available in the virtual world [6]. For the actions to take place, the user needs to be the main character of the environment rather than viewing the environment from another character's perspective. This helps avoid conflict between the other character and their own ego in the actions that take place. This conflict of actions occurs when the character makes an action that the user would not want to make, pulling them out of the virtual world [12]. Reactions of the environment should respond to the presence and actions of the participant [8] and there is an expectation of history of interaction [6] that the environment should acknowledge previous interactions.

These two elements of VR merge in the form of the mental model of the virtual world. Mental Models are the conceptual representation of ourselves and the world around us such as the way objects work, events happen, or people behave [6]. They are formed through experience, observation, training and instruction due to people's inclination to form explanations [13].

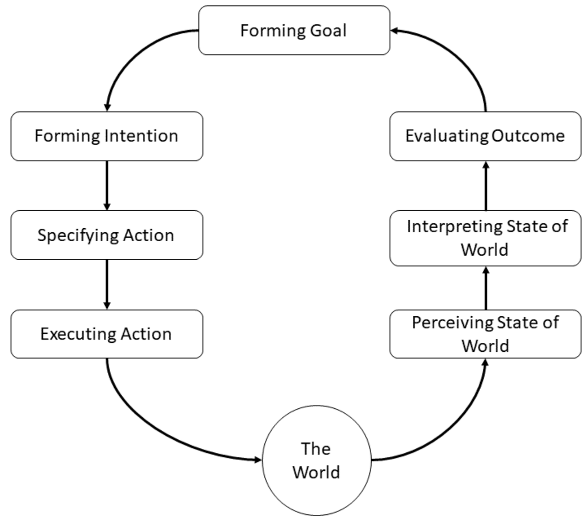

Figure 1: Action Cycle of Actions and Reactions

The action cycle, the continuing cycle of actions and reactions, is suggested to be a seven-staged, see Figure 1. It involves forming the goal and intention, specifying and executing the action, perceiving and interpreting the state of the world and evaluation of the outcome. The forming of intention and specifying the action is reliant on a person's mental model whereas interpreting the shape of the world and evaluating the outcome help reshape the person's mental model [6]. Essentially, what this means is that interactions help form our presence in a virtual world, creating an indirect link that interaction should increase the elicitation of emotion.

Hence, the environment needs to provide actions that fit the intentions of the participant and in a way that can be perceived and directly interpreted according to their expectations. These issues are referred to as the gulfs of execution and evaluation [13]. We can make use of Norman's seven design principles to make tasks easy for the participant to bridge these gulfs. These principles are as follows:

- Use both knowledge in the world and knowledge in the head – make use of mental models
- Simplify the structure of tasks.
- Make things visible: bridge the gulfs of Execution and Evaluation.
- Get the mappings right.
- Exploit the power of constraints, both natural and artificial.
- Design for error.
- When all else fails, standardize.

The design of the interactions in the environment attempt to follow these principles.

2.2 Eliciting Sadness

Throughout the study of emotion in VR, there have been many papers measuring emotion. The most prevalent emotions from these have been fear and anxiety [7, 11], but this paper is focused on sadness, and more specifically, "activating" sadness which relates to parasympathetic withdrawal and occurs when the person has agency where loss is imminent but is not inevitable [14, 15]. The main three factors which have influence over the elicitation of this emotion are attachment, empathy and events.

First, attachment has been shown to be a strong factor in sadness and grief with a strong correlation between the strength of an attachment bond and the extent of grief experienced [15]. This human attachment bond has also been shown to be equal to the attachment bond between a human and a pet. This means that in the context of VR, the use of avatars is not a hinderance [15, 16], but rather that these attachment bonds can be formed to be similar to human bonds. Interaction has an effect on attachment as bonds are formed from the dual elicitation of emotional responses from the participants and the shared activities between them [15].

Second, empathy influences both attachment and the elicitation of sadness [16, 17], and affective communication, the expression of feelings, is vital in forming this. In VR, for affective communication to occur, it is crucial that for any AI agents in the VE, the participant has agency belief, the belief about the virtual agent having agency. If the participant does not believe that a virtual agent acts and feels of its own volition, it is harder for them to feel empathy towards it.

Lastly, the elicitation of sadness is caused by events [18]. This event has several conditions which need to be met for the emotion to be evoked. First, the event is perceived as being inconsistent with the person's motives; Second, the participant's motive is to attain reward rather than to avoid punishment; Finally, it is event-directed – meaning that it was either caused by circumstance, no cause was specified, or there was a causal agent but that agency information was disregarded with focus on the event itself, the blame is not held on the agent. In terms of Virtual Reality, for sadness to be evoked, an event would need to occur in the environment which meets these three criteria. An example of this event could be a loved one passing away in a natural disaster, where there is no agent to be blamed and your motive is their safety, which is a reward and inconsistent with what occurs.

Two types of emotions can be distinguished in Virtual Reality [10]. The first is Fictional World emotions and arise from the illusion of being physically present in a fictional world – or virtual environment. The second is Artifact emotions and arise from the person's awareness that the fictional world or virtual environment is presented through an artifact, in this context it could be the awe someone feels when using VR technology for the first time. Fictional World emotion is the type required in our context, while Artifact emotions should be minimized to reduce noise.

2.3 Artificial Intelligence

An Artificial Intelligence Director is an AI that features a dynamic system for game dramatics and pacing, making changes to the game environment to enhance the experience [19]. It decides where game objects and agents spawn and how often these occur based on dynamic elements, such as the players skill, the time since a previous attack or how close they are to the end of the level. The AI Director is an overarching control of the system and makes changes to the overall experience.
On the other hand, AI agents make use of techniques such as decision trees, state machines, pathfinding, scripting and adaption to create a dynamic and improved experience. These are related to specific characters in the environment and focus on their behavior.

3 Design and Implementation

3.1 Approach

The goals for the virtual environment are formed from the goals and design constraints of the project. First, interaction needs to be provided in the environments as well as a way in which to manipulate the interactions that occur; Second, the environment needs to elicit sadness; Lastly, the environment needs to avoid inherent issues of VR, which are addressed later in this section, such as simulation sickness.

The environment was developed using the Unity game engine as the driving software engine, and the HTC Vive headset as the primary hardware, which has a resolution of 1080 x 1200 pixels per eye, a 90hz refresh rate, 110 angle field of view and two motion controllers that act as virtual hands. This produces an immersive experience in the developed environments as it resembles a closer similarity to reality than older generation headsets. The primary coding language in Unity is C#, which all the scripts, including the AI, interactions and interaction controllers as well as the driving narrative scripts were coded in. A small amount of code was written in a variation of HLSL to change the behaviors of existing shaders to achieve the desired effect.

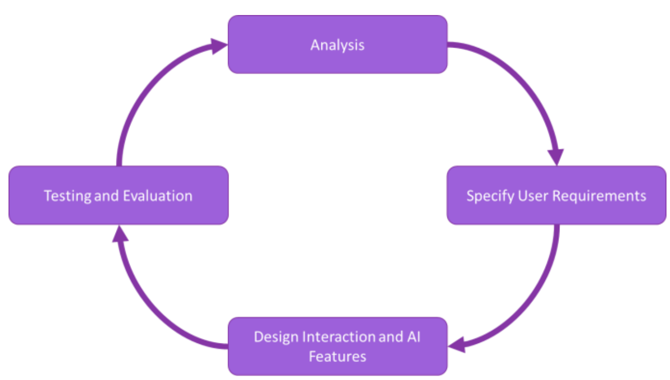

Figure 2: User-Centered Design Model

A software development project requires a design methodology that adequately prepares developers for design challenges and expectations. For development of the VE, the agile, user-centered design model was followed, which can be seen in Figure 2, which focuses on iterative software development cycles and constant user feedback for each cycle. At each iteration of development, through prototyping and demos, feedback was provided, and the environment modified based on this feedback. This feedback was provided from users in psychology and computer science fields. The features to be implemented for the next cycle would be outlined as well as any features that required modifications or fixes. Through these iterative cycles the levels of interaction were implemented, starting with the base control level and then adding features for the next level of interaction and so on.

Apart from the user-centered design model, the environments needed to accommodate and work around some inherent issues of VR. Firstly, any virtual environment that has movement or navigation must account for the possibility that users may experience simulation sickness, where they begin to feel nauseous or dizzy. This means that any navigation needs to have measures in place to avoid this. Teleportation, which will be described in section 3.3, was used with a fade in and out to cater for this. The fade helps avoid disorientation, while the teleporting is less likely to lead to sickness than other conventional movement. Secondly, uncanny valley effects, where an increase in realism remains appealing to the user only up to a certain point, and realism need to be balanced. The environment is designed to not achieve hyper-realism for this purpose, but still constructed in a way that is believable. Finally, a user's sense of presence that what is happening in the virtual reality is really happening is unlikely to recover once broken [8]. To account for this, the environment is designed to make the participant feel that events which occur that they have no direct control over are referring to them and their sensations. It also means that bugs or glitches that occur in the experience can break their sense of presence.

3.2 Previous Environment

The previous VE contained four scenes, namely: a hand selection scene, a pet store scene, a park scene and a vet scene. Each scene other than the first was used to convey different parts of the narrative and build up or cause the elicitation of different emotions. The user was invisible in the sense that they could not see any element or part of themselves throughout the experience. Each scene will be described, after which, the interactions and AI behaviors that were present will be outlined.

The hand selection scene is a simple scene where the predominant hand could be chosen, either left or right, allowing the participant to use that hand for the interactions, improving the presence of the experience. This scene is followed by the pet store scene, where the user would choose their puppy at the pet store. The scene consists of a room with props that resemble that of a pet store, such as food, toys, collars and three dog pens, each with a dalmatian puppy inside. While the scene is primarily in a room, there is a window which looks to an outdoor terrain which helps build realism. The puppy is chosen by the user opening the door to their pen. After the puppy is chosen, it approaches the participant and the scene ends.

The next scene is the park scene and is the most vital scene where the majority of the bonding and events take place. The puppy is now an adult dalmatian, with the model resembling the chosen puppy. The participant can interact with the dog by using a throw stick in their predominant hand to pickup and throw a tennis ball that the dog runs after and fetches. If left alone the dog either runs through some idle animations or walks to a specific spot on the ground and does a digging animation. The participant occasionally hears cars drive past. After an amount of time, the dog runs off barking towards the entrance of the park, where once out of sight, the sound of a car screeching and hitting something and the dog whining can be heard, followed by the end of the scene.

The final scene is the vet scene, where the participant appears in a room that resembles the reception of a vet. There are props such as pet food on one side, a chair and a receptionist counter on the other with a receptionist typing at a computer behind the desk. Behind the user is a window looking to an outdoor terrain. After a short wait, a vet in a white coat walks up with the tennis ball from the park in their hand. Once they stop by the participant, the vet says that they are sorry and that there was nothing that they could do to save the dog. The participant can then grab the ball and the experience ends.

The interactions in the scenes can be broken down as follows: In the pet store scene, the participant has the simple method of selecting the puppy as their pet, in the park scene they can use the throw stick to throw the ball, and in the vet scene they can take the tennis ball from the vet.

The AI from the previous environment was as follows: in the pet store, each puppy had a different set of animations that it would run through, and they had the same pathfinding script where they would walk towards the participant once chosen. The animation sets would give each a different sense of personality, one being calm, one being playful or high energy and one being friendly and excited. In the park scene, regardless of which one was chosen it would have the same behavior, and this behavior was controlled by decision trees. When the ball is thrown it barks and fetches the ball, when left alone it will select a random idle animation – either it would sit, scratch or go dig and after a certain time it would run to where it would be hit by the car.

3.3 Interaction

An important part of this project is the ability to manipulate the amount of interaction that occurs in the environment. This manipulation is achieved through the use of levels of interaction that are chosen at the start of the experience. There are three levels of interaction of which one is chosen when running the experience. The first level is a base or control level which contains the least amount of interactions, the second level has all the interactions from the first and a few additions and the final level has those same additions and more. In this section, each of the features implemented in these levels that are additions from the previous environment, as well as the intention of the interaction, will be described. A new simple scene was incorporated to allow the selection of the interaction level.

3.3.1 Level 0: Base Level. This level of interaction is intended to be used as a control and as such remains the closest to the previous environment. The first change is the addition of hand models. The hand models are the medium for interaction and allow the user to interact more naturally with objects throughout the different stages. The hand models have three animations, an idle animation, a grabbing animation and a pointing animation. When an object is grabbed, the grabbing animation occurs, otherwise the idle animation will play. The pointing animation will be explained in section 3.3.3. The hand models are also linked to colliders, objects in the Unity environment that ensure that they do not pass through other objects and thus preventing breaking of the realism of the experience. When these colliders touch an object, haptic feedback is provided to the controller through vibration, increasing the participant's sense that they are actually touching objects, and this is increased depending on the force applied. The hands allow you to pickup objects and throw them, applying velocity to the objects.

The second change that occurs is the removal of the selection of the puppy. Instead, there is only one dog present in the pet store scene and the door opens automatically after a short amount of time, after which the dog walks towards the participant. The player will still be able to feel that they are meeting their new pet and form a bond, but the agency of the interaction is removed as there is no longer a choice. This is done to create a greater contrast of interaction in the later levels.

The third change occurs in the park scene, where there is no longer a throw stick, but instead the participant can now use the hands and the behaviors that come with them.

The final change occurs in the vet scene where the vet no longer has the tennis ball, and instead the scene ends shortly after the vet finishes saying their line.

3.3.2 Level 1: Mid-Level. The hands continue to this level and most of the new changes are implemented in the park scene. Most of the interactions focus on either giving the participant a greater sense of agency or providing more interactions between them and the dog to attempt to strengthen the bond between them. The first difference to the previous stage is the puppy selection, allowing the participant to choose their pet from a selection of three dogs. This will give back the participant's agency in the pet store and increase the bond due to the feeling of choosing the one over the other two. This will also be the first time the dogs' different personalities will be apparent as they will behave differently. The participant will use their hand to open one of the dog pen doors to select the dog.

The ball being handed to the participant has also been re-added in the vet scene to try and increase the sense of loss that the participant experiences.

The rest of the changes occur in the park scene, firstly that the participant can now navigate through the park. This will be done through ray controls where they can teleport to a location where they place the arc reticule. This technique is taken over artificial locomotion through a trackpad as it avoids possible motion sickness, and this includes a short black fade to not disorient the participant. The navigation is added to give the participant more agency and allow them to approach their pet, explore or find new objects to interact with. This navigation also occurs in the pet store scene, allowing the participant to get close to each puppy and select any one of them.

This leads to the next addition, the ability to interact with multiple objects apart from the tennis ball. There are now new objects such as wooden cubes, blue bouncy balls, toy bones and a teddy bear, each of which can be picked up or thrown. The dog will react to the picked-up objects depending on their personality as well as the object picked up. This is done to create more interactions between the participant and the dog to increase the bond between them.

Lastly, the dog fetches objects as it explores. Depending on the dog's personality and the location it ends up at in the park, the dog may bring back the participant a fascinating item. This is either a toy bone it comes across or a teddy bear that it digs up. This is intended to create a special bonding moment and increase the bond between the pair.

3.3.3 Level 2: Advanced level. This level adds more interactive events which occur between the participant and their pet and allows more continuous interaction between them. It has all the interactions of the previous stages with some additions. Firstly, the participant is able to use hand gestures to communicate with the dog. The primary one is used to call the dog over in the park, but they are also able to get the dog to sit. The pet's response to the gestures depends on their personality. This gives the participant a sense of agency in controlling the dog as well as makes them feel that the dog is listening and reacting to them, which increases their bond. When holding the grip button on the controller, the pointing hand animation will play, after which if they complete the gesture and let go, the animation will stop, and the gesture will be recognized. The gesture recognition is implemented through the use of an existing software development kit, AirSig, that uses AI learning techniques to recognize 3D gestures.

The next interaction added is stroking. In the pet store as well as in the park, the participant is able to stroke the dog or puppy when they are nearby. The dog reacts to being stroked to give the participant a sense that the dog is experiencing it. This is one of the most common interactions people have with real pet dogs and is intended to increase the bond between the pair.

The third addition is an added butterfly interaction. When exploring, the dog will bark and run in a direction, waiting for the

participant if they are not close enough, looking excited in the direction they are headed. This will guide the participant to a small flower grove surrounded by butterflies. This location is intended to create a sense of beauty that the participant feels when seeing it, but the main reason is to increase the bond they feel at this shared experience with their pet.
The final interaction in the park occurs shortly after the accident. The participant is teleported next to their dog, which is lying on the ground breathing slowly. The participant can stroke the dog giving them the sense that they are there supporting their pet, increasing their bond and making them feel empathy before they experience the loss.
Lastly, in the vet scene the participant is handed the teddy bear instead of the tennis ball which they must grab from the vet before ending the experience. This brings back the companionship and bond they had formed to try and increase the loss they experience.

In addition to the stages, a new scene was introduced to indirectly enhance the interactions. Instead of a hand selection scene, a tutorial scene was implemented. This is to allow the participant to get familiar with the environment controls and use them comfortably during the experience instead of hindering it by having to learn, not being able to complete the interactions or not being aware that certain interactions exist. This makes the scenes as follows: Interaction Level Selection, Tutorial, Pet Store, Park and Vet Scenes. The tutorial shows the participant how to look around, snap the screen left or right (making looking around easier) and how to pick up or throw objects. If the mid-level, level 1, is selected, it also tells them how to teleport. If the advanced level, level 2, is chosen it also shows them how to stroke an object (a sphere in place of a dog) and how to gesture for a dog to approach them. After the tutorial instructions, the participant can change the hand models making the experience feel more personalized, changing to male or female hand models and a selection of three racial styles. This allows them to experience more presence in the environment as the hands match their own more similarly.

Figure 3: Tutorial Scene

Figure 4: Pet Store Scene


Figure 5: Park Scene


Figure 6: Vet Scene

3.4 Artificial Intelligence

The AI is implemented in two major ways, the first is through an AI Director that makes changes to the environment and overall narrative, based on the participants actions, and the second is the AI of the individual dogs. Both are implemented to try and create a more dynamic and realistic environment.

3.4.1 The AI Director. The AI Director primarily reacts to the participants actions, or more precisely, the lack of them. Based on the number of interactions or the different types of interactions that have occurred, it triggers events or behaviors in the dog to try and cause more interaction. Essentially, it makes changes to ensure that the pacing of the narrative is as desired and that a conducive bonding experience occurs. These changes are done through decision trees, an algorithm that only contains conditional control statements. Each change that the AI Director can make will be described and the reason for the change outlined.

Firstly, at a set time during the experience, if the participant has made too few interactions and not bonded enough with the pet, measured by the number of interactions that have occurred and the pet's current bond, the AI Director will force an increase in the pet's bond, this will be explained as part of the dog's AI in section 3.4.2. The increased bond causes the dog to behave more positively towards actions and this change tries to cause a shift in the dog's personality to skew the participant to take more actions with them. This change can occur in any level of interaction.

The next change which can also occur in any level of interaction, is that if the participant has not done any interactions with the dog in an amount of time, 20 seconds, the dog will run and pick up the tennis ball wherever it is and bring it back to the participant, making it seem like it wants to play. This change occurs to make the participant feel that they should play more with their pet and increasing the number of interactions that they make. The time was chosen to give the dog time to explore and return to idle animations without immediately picking up the ball but still acting regularly enough to have consistent interaction in the experience.
In the base interaction level, since the participant has no way of navigation, the AI Director will make the dog pickup the ball and bring it to the participant if it is getting too far away from them, without them having thrown it. Although unlikely this would prevent any interactions occurring.

The final changes that the AI Director can force, occur in the final interaction level and these exist to ensure that some of the added interactions of the stage occur. First, if the dog has not yet picked up the teddy bear well into the experience, it will do so, as this is vital for the vet scene. Second, if the dog has not yet led the participant to the butterfly interaction, it will do so. This once again corrects the narrative pacing and ensures that the bonding event occurs.

3.4.2 Dog Personality AI. The main changes in the dogs was giving them a sense of personality with differences in behavior between each as well as more realistic behavior. These changes are made to the AI of the dog primarily in the Park scene and the pathfinding of the dog is not touched upon as that is not changed. The first change to the existing AI was changing from only using decision trees to making use of a state machine, an abstract machine that can be in exactly one of a finite number of states at any given time and which based on the current state and given input, performs certain state transitions. In addition, the decisions within these states as well as the new transitions between them now make use of emotional behavior decision trees.

Machine learning techniques, such as Artificial Neural Networks (ANNs) and reinforcement learning were considered, but emotional behavior decision trees were chosen over these options. An ANN is an interconnected group of nodes, similar but far more basic than the network of neurons in a brain, and reinforcement learning is the process of training an agent using rewards and punishments. Both of these require large amounts of training data, that if done incorrectly will result in unrealistic behavior, and are better for complex systems that are difficult to model [20].

The emotional behavior decision trees are used for the decision making of the dogs and allows for each to have a different personality as it causes them each to have various reactions to certain events. It works the same way as a decision trees, but with the probabilities of the dog to take certain actions being dependent on specific "emotion" values of the pet. These values are set from the start, with different starting values for each of the three dogs but will be altered slightly through events or interactions between the pet and the participant. The "emotion" values that have been chosen are: *Curiosity*, *Playfulness*, *Obedience*, and *Bond* and each start at a value between 0.5 and 2, with a value less than one making it less likely to do certain activities, such as exploring with the curiosity value, and a value greater than one increasing the chances of the dog doing this activity. The choice of the first three have been chosen as they help direct what the pet would want to do, such as explore the park, how they react to a ball being thrown or how well they listen to the participant or stay by them after bringing the thrown object. These values are altered through some of the interactions, such as making the dog sit increases its obedience, making it more likely to listen when the participant gestures for them to return. Bond has been chosen to alter all interactions based on how the dog has been treated. Petting, playing and interacting with the dog increases this bond value. A dog with a higher bond is more likely to bring the participant to the butterfly interaction or listen more when gestured to come. This technique allows for the variation in personality as well as more dynamic behavior changes based on how the pet is treated. Each of these decisions occur in a certain state as a percentile chance, where a base percentage is multiplied by the sum of relevant emotion values, these values either affecting the outcome to be more likely or less likely. Each of the emotional behavior decisions will be outlined in the states that they occur.

1. *chanceSit = obedience * bond * 30*
2. *chanceExplore = 30 * curiosity*
3. *chancePlay = 25 * playfulness*
4. *chanceIgnoreGesture = (playfulness – (obedience * bond)) * 30 + (5 * object type)*
5. *chanceDisobedient = curiosity * 25 - (obedience * 25) * bond)*
6. *chanceLongExplore = curiosity * 30*
7. *chanceIgnoreThrow = (curiosity - (playfulness + obedience) * bond * 0.5) * 30 - (5 * object type)*
8. *chanceShowGrove = 35 * bond*
9. *chanceFetchToyBone - (playfulness * bond) * 45*

Figure 7: Emotional Behavior Decision Tree Formulae

There are nine states, which can be seen in Figure 8, that the dog is able to be in but may only be in one state at a time and this is the method that is used to decide the dog's current actions. There are variations for each state, and this is achieved through the use of the Emotional Behavior Decision Trees. Each state will be outlined with the actions that the dog takes when in the state and the occurrences of the emotional behavior tree decisions within it relating to the formula used to calculate its chance of occurring. Firstly, the Idle state is where the dog sits down if gestured (Figure 7.1), chases a ball if thrown, reacts if stroked, and otherwise plays around (Figure 7.3), sits down or scratches. It can also transition into the exploring state (Figure 7.2).

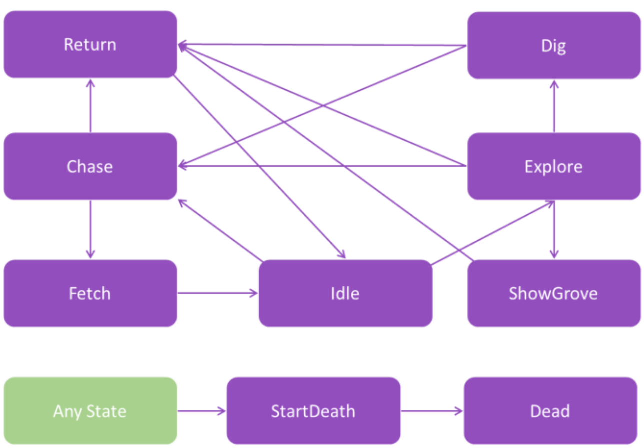

Figure 8: Dog State Diagram

When a ball is thrown from this state, the dog transitions to the Chase state, where it barks and runs after the ball. The participant can gesture to the dog to ignore the ball and return to them instead (Figure 7.4). This gesture has a chance of being ignored if the dog is playful and has a very positive attitude towards the object thrown. When the ball is picked up, the dog transitions to the Fetch state.

In the Fetch state the dog returns to the player with the ball, slowing down to a walk when nearby and stopping and dropping the object it is carrying when even closer. There is a chance that after dropping the object it will be disobedient and run off exploring instead of waiting for the participant to do something (Figure 7.5). If the dog is gestured to return before picking up the ball or while exploring, it will transition to the Return state which is similar to the Fetch state, except without an object.

When transitioning to the Explore state from the idle state, the dog will walk towards a random location, one of several preset locations. The time of its exploration is chosen to either be long or short (Figure 7.6) and whether it will ignore an object thrown by the participant while exploring is decided (Figure 7.7). The dog also decides whether it will show the participant the butterfly grove (Figure 7.8) and transition to the ShowGrove state. When the dog arrives at a digging spot, it transitions to the Dig state, otherwise after an amount of time has passed it transitions to the Return state. If it passes a nearby toy bone, it also has the chance of fetching it (Figure 7.9).

In the Dig state, the dog either switches to the Return state, or digs up the teddy bear and transitions to the Fetch state. It can be gestured to return while digging. The ShowGrove state controls the dog running towards the butterfly grove and waiting for the player to follow, before transitioning to the return state when both have arrived. The last two states are the DeathStart state and the Dead state, which first controls the dog running to the accident scene and secondly controlling it lying on the ground after the incident.

The combination of these states and the emotional behavior decision trees are what forms the emergent behavior of the dog in the park.

4 Test Design

4.1 Virtual Reality Heuristic Evaluation

The success of the project in creating an environment and AI that satisfy the aims, being able to manipulate the amount of interaction that occurs and to create an AI with dynamic and realistic behavior, is measured through heuristic evaluation done by experts in the VR field. Heuristic Evaluation [21] is a method for finding usability problems in a user interface design by having a small number of evaluators examine the interface against a set of usability principles, the heuristics. It has been shown that the aggregation of several evaluators to a single evaluation is able to do well in finding usability problems, even when the group consists of three to five evaluators [22]. It has also been shown that specialists in the field provide better evaluation than non-specialists [22].

In VR, heuristic evaluation is slightly different from traditional methods as they are extended to include VE-specific principles. VR can follow Sutcliffe and Gault's method of 12 specialized heuristics adapted for VR software [23] and the form used can be seen in appendix A. The structure of this measure revolves around three core categories:

1. Problems (from design classes)

The evaluator is required to present a problem within a design class that refers to a group of design features. For example, a design class would be interaction using virtual hands.

2. Associated heuristics

Once a problem has been identified, an associated heuristic is allocated to the problem. For example, interaction using virtual hands could present a problem with physics and will be associated with the Realistic feedback heuristic.

3. Severity of the problem

Finally, a severity rating from 0-4 is assigned to the problem and heuristic. Continuing the example above, a severity rating of 3 will be given to the problem, as unrealistic physics could break the sense of presence.

The evaluation was completed by three experts, which meets the minimum number of evaluators required for a valid heuristic evaluation.

4.2 Procedure

The evaluator was presented with a few open form questions both before and after the evaluation finding out their current emotional state, such as what emotions they were feeling, how strong these were, how comfortable they felt and what they were expecting in the environment. They were also asked for details about the VR setup that they would use for the evaluation. This was done to establish a base understanding for how and what the evaluator was feeling to give better context to the evaluation. Similar open form questions were asked afterwards, with the addition of which level of interaction they found most interesting, if they encountered performance issues, what they liked most about the environment and what element evoked the most emotion. While these questions do not affect the success factor of the project, they provide useful insight as to which interactions are the most effective.

```
====== Level Chosen: 2 ======
Tutorial Start: 40,21442
Tutorial End: 140,419
== Pet Store start: 144,1015
Puppy Stroked: 160,8153
Puppy Stroked: 165,4458
Puppy Stroked: 168,2773
Puppy Chosen: 169,71
== Park Start: 178,3266
Start Chase Ball: 185,779
Gesture - Sit: 194,7629
Dog Stroked: 207,3111
Dog Stroked: 210,3904
AI - Play: 216,3966
Start Chase Ball: 217,1621
AI: Gesture Obeyed: 220,6383
Dog Stroked: 227,4365
Start Chase Ball: 233,7879
Gesture - Sit: 246,9485
Gesture - Sit: 249,1129
Gesture - Sit: 256,8782
Dog Stroked: 260,0097
Gesture - Sit: 260,176
AI Director - Forced Get Teddy: 273,3358
AI: Start Exploring: 273,3686
Fetch Teddy: 290,4602
AI Director - Forced Grove Interaction: 305,8869
Grove Found: 319,5054
Start Chase Ball: 319,5373
Start Chase Ball: 337,5496
Start Running Away: 348,3567
Crash: 375,711
By Body: 380,7409
Injured & Stroked: 387,7262
Final Bond Value: 2,549999
Num Interactions: 19
== Vet Start: 389,9203
Teddy Used: 389,9203
Vet Speech: 411,3117
Handed Object: 417,8416
==== End of Experience ====
```

Figure 9: Example Log Showing Action That Occurred and Time

After completing the experience on each level of interaction, the evaluator completed the heuristic evaluation raising the issues with the environment. They were also able to make use of logs, as seen in Figure 9, that the environment generated showing what interactions occurred and when they occurred as well as decisions made by the AI Director and the Dog's AI.

5 Results and Discussion

5.1 Limitations

The primary limitation of this research was the inability to make use of user experimentation for the measurement of emotions. This was due to the COVID-19 pandemic and the health implications of these experiments making it unreasonable to measure the physiological changes, using measurement equipment required for skin conductance and heart rate, that participants would experience in the environment. Questions that relate to the emotional experience could also not be used, mainly due to the lack of availability of VR equipment, as the number of participants required for this is much larger than that which is required for a heuristic evaluation.

Lastly, the limited amount of time in which the project could be completed is a constraint that affects the user-centered design of the environment. An increase in the number of iterations before heuristic evaluation would result in better interactions in the environment and a decrease in heuristic issues as they would possibly be identified and resolved during development.

5.2 Heuristic Results

The heuristic evaluation was completed by three evaluators and they identified several problems that were apparent in the virtual environment. The problems that were identified by separate evaluators and were determined to be the same issue were grouped into a single problem. Due to two of the evaluators making use of a severity rating system of Low, Medium and High, this was the system used for all the evaluations. Where a number system was used, a rating of 1 was given a severity of Low, a rating of 2 or 3 was given a severity of Medium and a rating of 4 was given a severity of High.

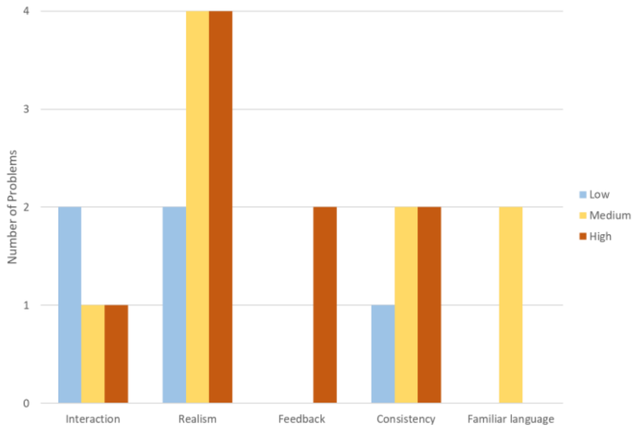

Figure 10: Problems for Heuristics by Severity

There was a total of 23 heuristic problems identified after aggregation, with realism having the highest number of problems and familiar language and feedback having the least number of problems. The problems that occurred in each heuristic will be discussed and the possible remedial actions will be outlined. It was also noted that a bug occurred in one of the tests that disrupted some of the further evaluation. This bug is believed to have caused many of the realism problems with differing setups as a potential cause, with differing hardware and versions of software causing unexpected behavior, and measures were put in place to address this afterwards.

5.2.1 Interaction. The first problem that occurred relating to the interaction heuristic with a low severity was the ability for the user to teleport outside of the park, going into areas that are not developed and breaking the sense of presence. There was also an issue with the user not being able to see the glass doors of the dog pens and not knowing how to open them. These could both be resolved easily, with boundaries for navigation being implemented in the park and frames being placed on the doors with clear instructions. The only medium severity problem that occurred was a lack in dog variance. This issue needs to be resolved in two ways, first with additional models to give visual variance between the dogs, and second with more variation in their AI. This problem requires additional assets to resolve and the lack of AI variation could mean that the dynamic behaviors do not make changes to a satisfactory extent. The high severity problem that occurred was that the gestures were not always working. This problem arises from two reasons, first that the instructions are not clear enough, and second is that the dog's behavior may lead it to ignore the gesture. Although this behavior is intended, it means that it may have a negative effect on the experience of the user instead of creating dynamic behavior. A way to resolve this issue would be to provide clear environmental feedback as to when the gesture has been recognized, allowing the user to know that the gesture worked but that the dog is ignoring it.

5.2.2 Realism. The low severity problems included grass textures, which occurred due to billboarding – a textured flat object which faces the camera - causing the grass to look 2-dimensional, and uncanny valley effects from models. These can be resolved with higher resolution assets. A lack of variation in the sound was a medium severity issue that can also be resolved with the addition of assets. The other medium severity problems were the lack of dog variety, which can be resolved as mentioned in section 6.2.1, the dog animations appearing buggy and the stroking animation seeming not natural. These latter two problems can both be resolved by improved use of existing animations for the models – both the dog and the hand – or by getting additional, more specific to the environment, animations designed which is time consuming and requires a high degree of skill. Of the four high severity problems, three of them occurred with the dog AI, first with it walking underneath the user, second with it running in place, and third with it picking up two balls at the same time. These issues are believed to have arisen from a bug due to differing setups and measures were placed in the state machine of the dog, making new conditions to ensure that the dog keeps a distance from the user and that it cannot pickup another ball when already carrying one. The final problem of high severity was an issue with the hand model, where it would rotate in a way that is different to the user's hand and grabbed objects would be out of sync with the hand. The rotation is likely due to the extended wrist of the hand model, which cannot mimic the rotation of a wrist making it a poor choice of asset. This is likely the reason behind the objects being out of sync of the hand.

5.2.3 Feedback. The issues that arise for this heuristic are both of high severity. The first is that the teleport delay is not shown to the user. This delay is explained in the tutorial but having a visual cue would lead the user to recognize why they are not able to teleport during the delay. The second is that the hand haptic feedback, that occurs through vibration, would sometimes be too intense when moving into an object. A way to fix this would be by clamping the maximum variation that touching an object can produce.

5.2.4 Consistency. Both one of the high and one of the medium severity problems that occurred arose from instructions not being clear or consistent throughout the experience. This issue could be resolved by either having written instructions in more obvious places or having voice instructions in each scene. An issue with having instructions outside of the tutorial scene is a possible break of presence, but having the participant not knowing how to do certain interactions could break the presence more severely. One problem noted was a transition being too sudden, which can be resolved by a fade. The selection of the hands had an issue where the background canvas would cover the options if the hand went too deep, and this could be resolved by either using ray casting or moving the background further away.

5.2.5 Familiar Language. Both these issues were of medium severity and are issues of clarity. First, the differences between the interaction levels is not clear, but can be resolved using a table instead of a list. Second, the tutorial instructions are not clear, and this can be resolved by more user testing, focusing specifically on the instructions and how to reconstruct them to be more understandable and clearer.

Figure 11: Table of Heuristic Issues Addressed Since Evaluation

Problem	Severity	Related Feature
User able to escape park	Low	Interaction
Pen doors not obviously visible	Low	Interaction
Hand selection screen goes blank	Medium	N/A
Haptic feedback too intense when moving into objects	High	Interaction
Dog moving under participant	High	AI
Dog picking up two balls	High	AI
Dog running in place	High	AI

From the heuristic problems identified, issues were fixed with focus on problems relating to the interaction and AI features that were implemented. These fixes were done by following the suggested remedial actions and to test if these suggested actions were viable and able to be done in a reasonable amount of time. There were no issues with any of the remedial actions and the problems were tested afterwards to ensure that they no longer occurred.

Apart from heuristic tests, the environment was able to log all changes that occurred in it, showing all the user actions and the changes that this caused to the dog's states. These were found to behave as expected, with the dog transitioning to the correct state. These logs also showed that the AI Director changes would make the dog change to the correct state.

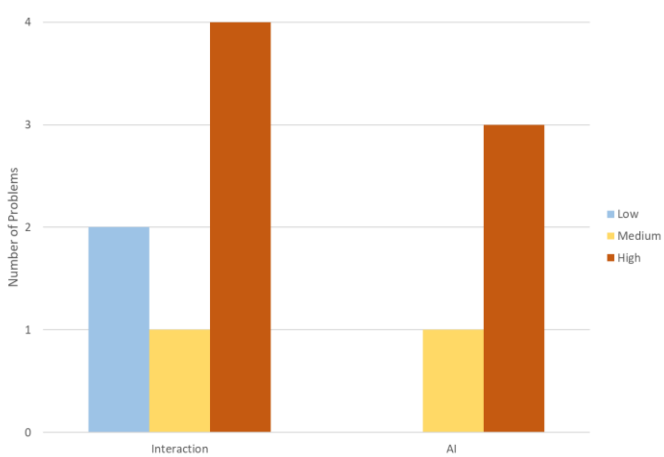

Figure 12: Problems for Heuristics for Features

5.3 Summary of Results

The heuristic problems can be categorized into those that either are caused by the interaction and AI features or those that are caused by other issues. Of the 23 problems, 11 can be put into this category, with 7 problems arising from interactions, drawn from the interaction, feedback and realism heuristics, and 4 problems arising from AI, drawn from the realism heuristics. Of the seven interaction issues, the two low severity issues and one of the high severity issues were resolved post-testing and all three of the high severity AI issues were resolved, as seen in Figure 11.

One of the issues with heuristic evaluation is that only problems are pointed out in the environment and can make it difficult to assess whether the environment is successful in achieving its goals. Additional comments made in the evaluation can help with this issue, and some of this feedback indicated that the environment worked as expected with the grabbing and throwing feeling good and natural for the evaluator. The interaction with the teddy bear was commented as being very effective in forming a connection and evoking emotion. The open questions indicated a clear increase in sadness after the experience, but due to this being done with only the heuristic evaluation, it cannot be used to determine the outcome.

Addressing the goals of the project, there were few issues remaining with the interaction features, primarily the hand being out of sync and feedback for teleporting, and none of these indicate that the levels between interactions have any heuristic issues. This lack of heuristic issues between levels means that the environment is able to manipulate the amount of interaction successfully. Most of the AI features were resolved, but the variation between the individual dogs was not adequate. The dynamic changes of the individual dog behaviors were not noticed by the user and thus did not impact the experience. This lack of effect means that, while these behaviors are believable, the environment is unsuccessful in creating dynamic behavior to the desired extent.

6 Conclusions

We sought to create an environment with the ability to manipulate the interactions that occur and to create an AI that behaves in a dynamic and believable way. Three experts in VR completed a heuristic evaluation where 23 problems were identified and given severity ratings, 7 of these relating to interactions and 4 of these relating to the dog AI. The interaction heuristics did not relate directly to the different levels of interaction but rather to how some of these interactions were carried out throughout the environment, meaning that the environment was successfully able to manipulate the interaction that can take place but that there are ways to make these interactions more natural for the participants. All of the high severity ratings that related to the dog's AI were resolved, meaning that the dog's behavior was believable. A variation in the dog's behavior was not noticeable to the participants and as such, the desired dynamic behavior, while present, had a less successful impact than desired in the environment. This is promising for the environment as it can now be used in future iterations focusing on interaction with a believable AI agent, but if a dynamic AI needs to be the focus of experimentation, more AI changes will need to be implemented.

There are several avenues which future work on this topic may lead. The initial, and most logical, next step would be to deal with the primary limitation of this project and perform user experimentation to measure the physiological changes that occur and to have them answer emotion evaluations. This would provide a more definite answer as to whether interaction results in an increase in the elicitation of sadness in VR. The second direction that could be taken would be the use of a variety of haptic devices and measuring the influence these cause with the interactions, such as gloves for better and more realistic control over the VR hands or a fluffy surface to mimic the dogs fur when being stroked. The final direction that is suggested is an investigation into which choices have a larger impact on the elicitation of emotion. These would be the choices with impact or consequence, such as choosing the dog, compared to playful or inconsequential interactions, such as throwing the ball or stroking the dog.

REFERENCES

[1] World Health Organization, 2017. Depression and Other Common Mental Disorders: Global Health Estimates. Geneva.

[2] D. Freeman, S. Reeve, A. Robinson, A. Ehlers, D. Clark, B. Spanlang, M. Slater, 2017. Virtual reality in the assessment, understanding, and treatment of mental health disorders. Psychological Medicine, Vol.47(14), pp.2393-2400

[3] Thomas D. Parsons, Albert A. Rizzo, 2007. Affective outcomes of virtual reality exposure therapy for anxiety and specific phobias: A meta-analysis. Journal of Behavior Therapy and Experimental Psychiatry, 2008, Vol.39(3), pp.250-261

[4] Matthew Price, Natasha Mehta, Eric B. Tone, Page L. Anderson, 2011. Does engagement with exposure yield better outcomes? Components of presence as a predictor of treatment response for virtual reality exposure therapy for social phobia. Journal of anxiety disorders, 25(6), pp.763-770.

[5] Amy Karon, 2015. Exposure-based therapy best for complicated grief. Clinical Psychiatry News, vol. 43, no. 2, pp. 25.

[6] Peter van der Straaten, Martijn Schuemie, 2000. Interaction affecting the sense of presence in virtual reality. Delft University of Technology, Faculty of Information Technology and System, 67.

[7] Julia Diemer, Georg W. Alpers, Henrik M. Peperkorn, Youssef Shiban, Andreas Mühlberger, 2015. The impact of perception and presence on emotional reactions: a review of research in virtual reality. Frontiers in psychology, 6, Article 26.

[8] Mel Slater, 2009. Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments. Philosophical Transactions: Biological Sciences, Vol. 364, No. 1535, Computation ofEmotions in Man and Machines (Dec. 12, 2009), pp. 3549-3557

[9] Joschka Mütterlein, 2018. The three pillars of virtual reality? Investigating the roles of immersion, presence, and interactivity. In Proceedings of the 51st Hawaii International Conference on System Sciences.

[10] Valentijn T. Visch, Ed S. Tan, Dylan Molenaar 2010. The emotional and cognitive effect of immersion in film viewing, COGNITION AND EMOTION, 24:8, 1439-1445, DOI: 10.1080/02699930903498186

[11] Jaehee Cho, Tsung-Han Lee, Joel Ogden, Amy Stewart, Tsung-Yu Tsai, Junwen Chen, and Ralph Vituccio. 2016. Imago: presence and emotion in virtual reality. In ACM SIGGRAPH 2016 VR Village (SIGGRAPH '16). Association for Computing Machinery, New York, NY, USA, Article 6, 1–2. DOI: https://doi.org/10.1145/2929490.2931000

[12] Eric Darnell, Michael Hutchinson. 2016. Invasion!: exploring empathy and agency in VR. In ACM SIGGRAPH 2016 VR Village (SIGGRAPH '16). Association for Computing Machinery, New York, NY, USA, Article 10, 1–2. DOI: https://doi.org/10.1145/2929490.2929494

[13] Don Norman, 1988. The Design of Everyday Things (2002 ed.), Basic Books.

[14] Mariko Shirai, Naoto Suzuki, 2017. Is Sadness Only One Emotion? Psychological and Physiological Responses to Sadness Induced by Two Different Situations: "Loss of Someone" and "Failure to Achieve a Goal". Frontiers in Psychology, 8. doi:10.3389/fpsyg.2017.00288

[15] Nigel P. Field, Lisa Orsini, Roni Gavish, Wendy Packman, 2009. Role of Attachment in Response to Pet Loss. Death Studies, 18 March 2009, Vol.33(4), pp.334-355.

[16] Takashi Numata, Hiroki Sato, Yasuhiro Asa, *et al*., 2020. Achieving affective human–virtual agent communication by enabling virtual agents to imitate positive expressions. Scientific reports, Vol.10(1), pp.5977

[17] Felicia R. Baltes, Andrei C. Miu, 2014. Emotions during live music performance: Links with individual differences in empathy, visual imagery, and mood. Psychomusicology: Music, Mind, and Brain, 24(1), 58–65. https://doi.org/10.1037/pmu0000030

[18] Ira J. Roseman, Martin S. Spindel, Paul E. Jose, 1990. Appraisals of Emotion-Eliciting Events: Testing a Theory of Discrete Emotions. Journal of Personality and Social Psychology, 1990, Vol.59(5), pp.899-915.

[19] Michael Ambinder. Biofeedback in Gameplay: How Valve Measures Physiology to Enhance Gaming Experience. In Game Developers Conference, 2011.

[20] Darryl Charles, Stephen McGlinchey, 2004 The past, present and future of artificial neural networks in digital games. In Mehdi Q, Gough N, Natkin S, Al-Dabass D (eds) Proceedings of the 5th international conference on computer games: artificial intelligence, design and education. The University of Wolverhampton, pp 163–169.

[21] Jakob Nielsen, Rolf Molich, 1990. Heuristic evaluation of user interfaces. CHI '90: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 1990, pp.249-256.

[22] Jakob Nielsen, 1992. Finding usability problems through heuristic evaluation.. CHI '92: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 1992, pp.373-380.

[23] Alistair Sutcliffe, Brian Gault, 2004. Heuristic evaluation of virtual reality applications. Interacting with Computers, Volume 16, Issue 4, 2004, pp. 831–849.

Appendix A

Heuristic Evaluation Form

Virtual Environment (fear or sadness): SADNESS

Instructions: Use the table below to write problems associated with the elements in the Virtual Environment. For each problem, determine which heuristic relates best to the problem, and write a corresponding severity rating. You may add additional elements and add as many problems for each event as you like.

HEURISTICS	SEVERITY RATINGS
1 – Interaction	0 – No usability issue
2 – Agency of user (freedom and sense of control)	1 – Slight usability issue, needs not be fixed unless there is extra time available
3 – Simulation sickness (comfortability)	2 – Minor usability issue, low priority
4 – Realism	3 – Major usability issue, important to fix
5 – Familiar language	4 – Extreme usability issue, needs to be fixed immediately.
6 – Consistency	
7 – Recognition over recall	
8 – Flexibility and Efficiency	
9 – Aesthetic and simplicity	
A – Additional heuristic (please specify)	

EXAMPLE ELEMENT			
PROBLEM	**HEURISTIC**	**SEVERITY**	**NOTES**