

The Use of Emotions to create Believable Agents in a Virtual Environment¹

Karthis Selvarajah
Computing Department
Macquarie University
North Ryde, 2109

Debbie Richards
Human Centred Systems Group
Macquarie University
North Ryde, 2109
+61 2 9850 0567
richards@ics.mq.edu.au

ABSTRACT

In the past emotions have been dismissed as a distraction to the logical, scientific thought process. More recently however, the importance of emotion in human-like intelligence and behaviour has been identified. This project aims at exploring this aspect of Artificial Intelligence by modeling the ability to display emotions in autonomous software agents within the constraints of a virtual environment. The motivation behind this is to determine whether the behaviour of these agents will cause the human participant to interact with the agent as if interacting with other humans. We have created an Agent-Cocktail Party World for this purpose and an Avatar-Cocktail Party World for the purpose of studying the psychological phenomenon of Ostracism. Our results show that the addition of an emotion-based intelligent component was able to make a statistically significant difference to the experimental condition by creating a more realistic environment in which to simulate the Punitive Ostracism condition.

Categories and Subject Descriptors

I.6.7 [Computing Methodologies] Simulation and Support Systems, *Environments*, J.4 [Computer Applications] Social And Behavioral Sciences, *Psychology*.

General Terms

Design, Experimentation, Security, Human Factors.

Keywords

emotion-based agent architecture, ostracism, psychology, virtual reality.

1. INTRODUCTION

Many scientists conducting research into emotions have understood that emotions play a significant role in cognitive processes. It is their influence on these cognitive processes that characterize emotions, and the ability to display emotional expression, as being a fundamental part of realistic human-like behaviour. Scientists (eg.[4, 12]) have developed software systems that model emotional phenomena in software agents. Their projects have been tested using computers only, not with a

human participant to determine if their models are realistic enough, to make a human participant believe they are interacting with another human being. This project aims at exploring this aspect of Artificial Intelligence by modeling the ability to display emotions in autonomous software agents within the constraints of a virtual environment. The motivation behind this is to determine whether the behaviour of these agents will cause the human participant to interact with the agent as if interacting with other humans. The Agent-Cocktail Party World was designed for this purpose.

The Cocktail Party World is a virtual environment that was initially proposed by psychology Professor Kipling D. Williams for the purpose of furthering his study of Oblivious Ostracism. The environment we created for Williams, termed the Avatar Cocktail Party World, contained scripted avatars (graphical representations of human beings) rather than artificially intelligent software agents. We have developed an additional virtual environment using agents in order to determine whether the emotion enriched software agents firstly, caused the human participant to perceive the Agent Cocktail Party World as being more realistic than the Avatar Cocktail Party World. Secondly, but outside the scope of this paper, to show that psychological experimentation and testing could be enhanced by techniques from artificial intelligence, as the participants react more naturally to autonomous software agents than to the scripted avatars thus enhancing the validity of the study.

In the next section we define some key aspects of our study. In Section 3 the emotion-based architecture is briefly introduced. Section 4 describes our experiments. Conclusions can be found in the final section.

2. BACKGROUND

In this section we define the key components within our study. These components are: ostracism and its forms; immersive virtual environments; the Cocktail Party World and; modeling emotions in an artificial intelligent agent.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee
AAMAS'05, July 25-29, 2005, Utrecht, Netherlands
ACM 1-59593-094-9/05/0007

2.1 Ostracism

Ostracism is a powerful social phenomenon that has an impact on the lives of many people on a daily basis. To date, Ostracism and its profound emotional and social consequences have not been thoroughly studied. Interest in understanding ostracism has increased since the shootings at Columbine High School revealed that the teenage shooters acted on their feelings of isolation and exclusion.

The potency of Ostracism lies in its simplicity. Unlike other aversive interpersonal behaviour that requires explicit action from the source, Ostracism is just as powerful; with the target left to infer what has brought on the Ostracism. It is precisely this ambiguity from the target's perspective that causes the four fundamental needs of belonging, self-esteem meaningful existence and control to be threatened. Williams [14] has defined three types of Ostracism: **Punitive Ostracism** where a group or individual deliberately ostracizes an individual as a form of punishment by using behaviors like the silent treatment or excommunication; **Defensive Ostracism** when a group or individual deliberately ostracize a target out of fear and **Oblivious Ostracism** where the target is unintentionally ostracized and made to feel as if they are inconsequential, invisible and as if they don't exist. This form is the most difficult to study as it is difficult to simulate situations where one may feel obliviously ostracized alone, without any punitive connotation. As a result, only few accounts about Oblivious Ostracism exist. However Williams hypothesizes that Oblivious Ostracism maybe the most potent form of Ostracism because of its profound impact on the target.

2.2 The Cocktail Party World

The Cocktail Party World (CPW) is an Ostracism paradigm proposed by Williams and formally documented via our work. The Cocktail Party World paradigm aims to create a situation where one feels Obliviously Ostracized. In the Oblivious Ostracism condition of the Cocktail Party World, the human participant is made to feel invisible and non-existent while they attend the cocktail party. This is done by not having the other guests in the Cocktail Party World acknowledge the presence of the participant.

In order to capture the invisibility which is characteristic of Oblivious Ostracism, Professor Kipling D Williams wanted to have one of the guests at the Cocktail Party walk right through the human participant. The participant, therefore, is left feeling like a silent observer, as if they are really not present at the party. It is for this reason that Immersive Virtual Environment Technology was chosen as the medium to create the Cocktail Party World experiments, as it is impossible to simulate this condition using human actors in a laboratory.

The Cocktail Party World is a simulation of a cocktail party with many guests in a room, drinking and non-verbally interacting with each other. The room that the virtual cocktail party is held in is similar to a formal room of a hotel or convention centre. The guests in the room are in groups or clusters, talking with each other. The guests move their mouth as if talking, but the participant will hear nothing. Background music and ambient party sounds are used to give the illusion that the people in the room are in fact, conversing with each other. The participant should assume that the guests do not know each other. The

participants can see the guests in the room walking around the room, to the bar or to the other groups.

The guests in the room are between 18 – 25 years old, close to the age of the participants. The guests are initially of mixed race. Future work will consider modifying race, gender and clothing to increase feelings of alienation.

2.3 Virtual Environments

Virtual Environments (VEs), or more commonly Virtual Reality (VR), can create the illusion of a real world by providing synthetic sensory information to the sight, sound and touch senses. The brain interprets this synthetic information as if received from the environment physically surrounding a person. As a result, the person experiences the psychological phenomenon known as presence, where the person is made to feel as if they are immersed in the Virtual World. These worlds are termed 'Immersive Virtual Environments' (IVEs) [11].

Our Immersive Virtual Environment required a graphical display, termed the Head-Mounted Display (HMD) to view the virtual world [3]. Speakers were used for the participants to hear sounds in the world. The presence created within the virtual world must be seamless in order to replicate the real world as closely as possible. A computer capable of rendering (drawing) many frames per second was required. To interact with the world motion-capture body suits or data gloves are required, but were not used in our study.

In the Cocktail Party World experiments, it is important to control what the user experiences, as we are trying to simulate a very intricate condition – the feelings associated with Oblivious Ostracism. Using Immersive Virtual Environment Technology (IVET) we as the investigators can control what the avatars in the VE do, how they interact with each other and most importantly how they interact with the human participant.

Increasing the number of avatars/agent in the CPW reduces the update rate which can cause discomfort for the participant, creating feelings of nausea and dizziness [3,11]. Thus, the program was developed to run at a rate which does not cause such discomfort to the participant.

2.4 Modelling emotions in intelligent agents

For the human participant to interact with the avatars in a manner that is similar to interacting with other humans, we need to model realistic looking avatars that display some form of human-like intelligence. Within an IVE, it is not actual but perceived intelligence that is significant: "Philosophers and psychologists have argued for over a century that what matters in human-human interaction is the individual's subjective beliefs about each other, not the objective truth of the interaction." [1].

The question of what initiates the perception of an avatar's intelligence arises from this subjectivity. Bailenson [1] suggest that "there are many likely variables [that affect this subjectivity and] interact in complex ways: photo-realism, non-verbal behaviour, autonomy and interactivity". Bailenson's research found that humans respond to virtual avatars, even those avatars that lack photo-realism, in ways very similar to their response to other humans.

However, one problem that was identified with early constructions of avatars was their stiff behaviour. Designers soon came to the conclusion that the avatars were “lacking emotions” [5]. It was identified that non-verbal behaviour is extremely significant for an avatar to act in a realistic manner. This non-verbal behaviour includes facial expressions, raising eyebrows, the movement of the head and mutual gaze [5].

Researchers have found that agents, who exhibited some emotional-type awareness, were treated like real human beings [1]. In a study conducted to measure interpersonal distance between a human participant and a human avatar, the human participant maintained a much larger interpersonal distance, similar to a real human, when the avatar maintained mutual gaze or smiled [1]. This suggests that a more realistic human avatar interaction response can be derived if the avatar displayed some form of emotional awareness, as emotions are a distinct characteristic of human like intelligence [4, 13].

While conducting research at Disney Studios Bates [2] discovered one characteristic that is essential for modeling realistic human behaviour; appropriately timed and clearly expressed emotions in characters. Disney animators were successfully able to “create the illusion” of human intelligence and life as opposed to ‘genuine life’ which would suggest that the entities really are ‘alive’ instead of just seeming to be so, by animating facial expressions in their characters. Revealing that having the capability to display emotions is a clear sign of human-like behaviour.

Although many different aspects comprise ones emotional state (such as tone of voice, volume, body and hand gestures) the most effective method of depicting one’s current emotional state is through facial expressions. Thus, for the avatars and agents in the Cocktail Party World, we chose to use the seven universal emotional expressions identified by psychologist Paul Ekman [7,8] as indicators of the emotions being simulated by our avatars and agents. The difference being that the avatars display these expressions when programmed to do so, the agents display these expressions autonomously as they switch between different emotional states. To model the emotional transition we required an agent architecture.

3. EMOTION-BASED ARCHITECTURE

Humans are constantly collecting emotional information from their interactions with others [6-9]. The information collected can be as simple as the person’s mood or attitude or can be more detailed information about a person’s current emotional state. It is this interaction that is important to our study. We would like to determine whether Virtual Environments that allow the human participant to interact with another (computer simulated or real) person emotionally, increases the realism of the environment.

Our architecture is based on key concepts from the work of Oliveira and Sarmento [12]: namely the process of emotional elicitation and the emotional accumulator. Emotional elicitation is the process that stimulates and triggers particular emotional phenomena. Emotional Accumulation describes how conscious the agent is about experiencing the emotion. Furthermore Oliveria and Sarmento identified three types of emotional phenomena; specific emotions, moods and dispositions. For this study we chose to model only specific emotions, as it was the effects of short-term ostracism that was of interest and moods and emotional

dispositions are prevalent for a long period of time. The specific emotions that will be used are those identified by psychologist Paul Ekman [8,9] as being globally identifiable. These emotions are anger, surprise, fear, disgust, happiness, sadness and the neutral expression.

Oliveira & Sarmento identified that the elicitation process involves the agent assessing its capability to achieve a certain goal given the environment and the agent’s internal state. In the cocktail Party World, the environment refers to hotel/conference room where the cocktail party is being held. This definition of the environment was revised to the other twenty agents in the room. This is because otherwise the room contains only inanimate objects, whose state in the Vizard software cannot be determined.

The term ‘goal’ is an explicit formulation about desirable future states that the agent seeks to achieve, as well as the implicit states that must be reached or maintained in the process. The goal in the Cocktail party World is to make a friend. To accomplish this goal the agents in the room simply interact with other agents in the within their vicinity. The varying degrees of success in the agents attempt to achieve the goal, is depicted by the emotional expressions on the face of the agent.

Initially each agent is assigned one random emotion. This is because at a real cocktail party, each guest arrives at the party with an emotional state that is (generally) not related to the emotional state of any other guest at the party.

Each agent in the environment is capable of interacting with other agents in a bidirectional manner within its vicinity. The term vicinity describes something similar to human personal space. The vicinity region is approximately a circular region of radius half a meter in room dimensions. The vicinities of agents overlap when the agents come together and this further promotes the two-way interaction between agents. The agent receives feedback from the environment in terms of a new emotional state, which is calculated, based on the collective emotions of the other agents and the human participant in its vicinity.

The algorithm we have developed assigns a weight to each emotion. The weights are initially assigned to each emotion by classifying the emotion as either positive or negative please refer to Table 1. Of the seven basic emotions, neutral is neither positive nor negative, and we have two positive and four negative emotions.

	Emotion	Assigned Weight
Positive Emotions	Happiness	6
	Surprise	4
Neither	Neutral	0
Negative Emotions	Sadness	-1
	Fear	-2
	Anger	-3
	Disgust	-4

Table 1: The emotions and associated weights

Once these emotions were ranked in order of intensity this the weights were assigned to ensure that the total score of positive emotions inversely equaled the total score of negative emotions.

The agent calculates its next emotional state based on the average emotion of each member of the group and the human participant if within the vicinity. The participant is asked to choose which of the seven emotions best describes their current mood. Their choice is included into formula (1) The average is then computed (formula 2) and then added to the agent's current emotional state to determine the next emotion (formula 3). Formally,

Let $E_{i,j}$ denote the current emotion of member i in group j

Let n_j denote the number of members of group j .

Let S_j denote the sum of the emotions of group j

Let A_j denote the average emotion of group j

Let E_{ij}' denote the next emotion of member i in group j

Then,

$$S_j = \sum_{i=1}^{n_j} E_{i,j} \quad (1)$$

$$A_j = S_j / n_j \quad (2)$$

$$E_{ij}' = E_{ij} + A_j \quad (3)$$

Changing formula 3 to an average would be more intuitive but we found in the pre-pilot test that the emotions and thus faces changed too gradually for the participants to notice the change. To avoid emotions being trapped at each minimum or maximum point, there is a lower and upper bound of 15 and -15. When these limits are reached, that agent's emotion value is set to neutral.

4. THE CPW EXPERIMENTS

Two CPWs were created, one which included avatars and another which included agents. The Avatar Cocktail Party World experiments were conducted for the purpose of Professor Kipling D. William's research into Oblivious Ostracism. Each avatar within the Avatar Cocktail Party World is completely programmed to behave in a certain way. The avatars are only capable of displaying the seven basic human facial expressions mentioned above. The avatars will display the angry and disgusted expressions when simulating the Punitive Ostracism condition. The fearful and surprised expressions will be displayed when simulating the Defensive Ostracism condition. The happy expressions will be displayed when simulating the Inclusion condition.

Although an avatar has the capability to display expressions the avatar has no knowledge of either its current emotional state, the emotional states of the other avatars around it or the participants emotional state. This is how the avatar differs from the agents within the Agent Cocktail Party World. This reflects the aims of the Avatar CPW (Aim1: to collect information regarding the subjective experience of being Obliviously Ostracized, Aim2: To show that Oblivious Ostracism is the most potent form of Ostracism) are psychology based issues and outside the scope of this paper.

The Agent Cocktail Party World is identical to the Avatar Cocktail Party World except that the guests within the Agent CPW are agents. Each agent has the capability to display a range of facial expressions, to have knowledge about its own emotional states, the emotional states of the agents around it and also the emotional state of the human participant. Each agent will be

constantly displaying a facial expression, which reflects the agent's current emotional state. The emotional state is updated periodically every 30 seconds. Although most specific emotions take effect for a period of a few seconds to minutes, we chose to update the emotions every 30 seconds because the agents are using a graphical means to depict their expressions, and it would be odd for an agent to display the same expression on their face for a period longer than a few seconds.

Additionally the agent's behaviour is modified depending on the world and the form of ostracism the participant has been assigned to. In the agent world the agents express its current emotional state on its face. The agents react to the participant in different ways based on the type of Ostracism.

In the punitive condition, it is only when the participant approaches the agent does the agent turn around and display the angry or disgusted expression. However, as soon as the agent turns back to the group, it displays its current emotion. For example, when the agent turns back to the group it may start smiling so that the ostracism is seen as deliberate. This is similar for the defensive and inclusion conditions.

For the punitive and defensive conditions, when the avatars turn around to face the participant and display the relevant expression, they do this for a period of 4 seconds, then turn back and face their group. In the inclusion condition, the agent turns around, and remains smiling at the participant until the participant leaves the agent's vicinity area. The aim of the Agent CPW was: To determine whether the addition of artificially intelligent agents displaying emotions increases the realism of the agents, which increases the realism of the Virtual Environment. It was expected that the increased realism of the agent will cause the realism of the environment to increase.

As part of the study the following seven research questions were investigated:

RQ 1: Does the order of the worlds affect the results?

RQ 2: Does artificial intelligence increase the potency of the ostracism?

RQ 3: Is the second world always worse?

RQ4: Does the participant accurately recognize the emotions?

RQ 5: Does the participant perceive the ostracism?

RQ 6: Does the Agent Cocktail Party world seem more realistic?

RQ 7: Is the participant able to perceive the reasons for their ostracism clearly?

Research Questions 2, 4 and 6 are of greatest relevance to this paper as they are directly concerned with the addition of emotion to the Agent CPW.

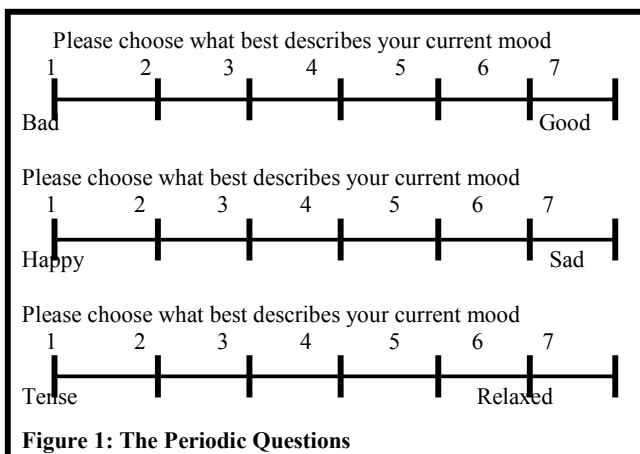
4.1 Experimental Design

As we were interested in testing for Oblivious, Punitive and Defensive Ostracism using both avatars and agents, the Cocktail Party World experiments have four different conditions that can be assigned to participants: Inclusion, Oblivious, Punitive and

Defensive. The subjects were further divided into whether they were given the Avatar Cocktail Party World first or the Agent Cocktail Party World first. Each participant was assigned a condition and received the Agent Cocktail Party World for that condition and the Avatar Cocktail Party World for that condition in a random order. For example, subject A might have received the Agent Oblivious Ostracism World followed by the Avatar Oblivious Ostracism World.

Each participant followed these steps:

1. Complete Facial Expression Identification Test. This is the first virtual environment the participant experiences. In this test, a male avatar from the Cocktail Party World displays the six emotional expressions one at a time. The purpose of this test is to ensure that subjects are able to recognize the emotions claimed by Ekman to be universally recognizable.
2. Experience the first world. During both the Agent and Avatar Cocktail Party Worlds there were questions that appeared in the HMD periodically every minute. The questions are asked five times within each world. The questions and the relevant scale are presented in Figure 1. This data was collected so that Williams could determine whether and to what extent the feeling and type of Ostracism was developing in the participant. For the Agent world, this data was also used to compute the next emotion of each agent in the participant's vicinity .
3. Complete the paper based survey (Figure 2) while the second world is being configured.
4. Experience the second world, periodically answering the questions in Figure 1 as in step 2.
5. Again complete the paper based survey as shown in Figure 2 while the alternate world is being configured.
6. Complete the World Comparison Survey: Each participant was required to answer a number of questions comparing World 1 and World 2 including which world they thought was more engaging.
7. Complete the Technology Survey regarding their experience with the IVET.
8. Figures 3-6 provide screenshots of the Oblivious, Punitive, Defensive, and Inclusion Avatar Experiments.



- Q1: During the Cocktail Party, I felt good about myself
- Q2: I felt that the other guests perceived me as a worthy person
- Q3: I felt that the other guests perceived me as a likeable person
- Q4: I felt accepted by the other guests
- Q5: I felt as though I made a connection and bonded with the guests at the Cocktail party
- Q6: I felt like an outsider during the cocktail Party
- Q7: I felt non-existent during the Cocktail Party
- Q8: I felt in-control during the Cocktail Party
- Q9: I felt as though my existence was meaningless during the Cocktail Party
- Q10: I felt ignored and excluded by the guests at the Cocktail Party
- Q11: I felt rejected by the other guests at the Cocktail Party
- Q12: I felt include by the other guests at the Cocktail Party
- Q13: I felt invisible to the other guests at the Cocktail Party
- Q14: I felt like I was being punished by the other guests at the Cocktail Party
- Q15: I felt like I scared the other guests at the Cocktail Party
- Q16: Do you think this room contained guests controlled by Computers?
- Q17: Do you think this room contained guests controlled by actual humans?

Figure 2: The Survey Questions



Figure 3: A screenshot of the Avatar-Oblivious Experiment

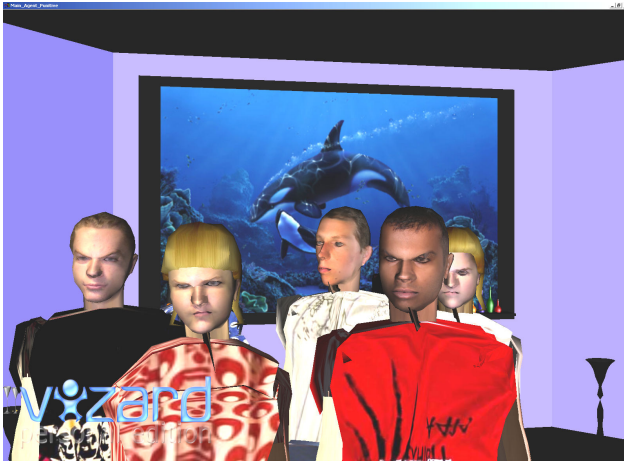


Figure 4: A screenshot of the Avatar-Punitive Experiment



Figure 5: A screenshot of the Avatar-Defensive Experiment



Figure 6: A screenshot of the Avatar-Inclusion Experiment

4.2 Results

In total, 54 participants, volunteered for the study. Two participants were long-sighted and could not read the text on the HMD without their glasses and thus could not participate in the experiments. An additional three participants could not complete the experiments as they became dizzy during the first world. In the end, 49 participants completed the experiments.

The 49 participants were divided into three groups of 12 and one group of 13. Each group of 12 participants participated in either one of the Oblivious, Punitive or Defensive Studies. With 6 participants in each group experiencing the Avatar Cocktail Party World as their first world, followed by the Agent Cocktail Party World. The remaining 6 participants in the group experienced the Agent Cocktail Party World first followed by the corresponding Avatar Cocktail Party World.

The last group of 13 participants experienced the Null condition, which was the inclusion condition. Seven participants experienced the Avatar Cocktail Party World first, followed by the Agent Cocktail Party World.

The expression ratings from the Facial Expression Identification Test were in keeping with Paul Ekman's findings (Note: The exact percentage values were not published in Ekman's paper). The emotions mostly easily identified were happiness and anger. The emotions Disgust and Sadness were identified by half of the participants, and finally the emotions Fear and Surprise were identified correctly only very rarely. The impact of this pre-test result will be discussed.

In order to compare the differences between the Agent world and the Avatar world, the differences between the means of the rejection index of each condition (Oblivious, Punitive, Defensive and Inclusion) need to be compared. A simple T-test was done to compare the differences in the means calculated using an ANOVA test. For the Punitive Ostracism we were able to reject the null hypothesis which stated "that the mean values of the avatar and agent groups are the same" and conclude with a confidence level greater than 99% that the participants were able to detect a difference with the Artificial Intelligence component and that it impacted their interpretation of the environment, making them feel worse in the Punitive Ostracism condition. This result confirms what Paul Ekman [8] suggested and what Bates [2] discovered with Disney animators, that humans actually are able to interpret and react to emotions being displayed to them in interaction even with a computer-simulated object. By the same token, it can be seen that the other conditions (Oblivious, Defensive and Inclusion) were invalidated because of this.

For the Defensive Ostracism condition, the emotion fear and surprise are the two primary stimuli. Paul Ekman, in his research found that most people confused these expressions with each other. For our Defensive Ostracism study if people had confused fear and surprise with each other would not have been a problem, as these two expressions are the only expressions being displayed primarily in this condition. However, participants did not confuse these expressions with each other, but two different emotions completely with 68% of the participants confusing fear with sadness and 51% of participants confusing surprise with the neutral expression. As a result we had to conclude that for over 50% of the Defensive Ostracism cases, the study became

invalidated, as the participants were unable to properly identify and comprehend the stimulus they received.

For the Inclusion condition in both the Agent and the Avatar worlds the mean scores are similar to the mean scores for the Oblivious condition, which suggests that participants felt ignored even in the Inclusion condition. While the happy emotion had not been confused we believe that because there were on 2/7 positive emotions and only one (happy) being modeled there was a predominance of negative feeling in the worlds. Finding more positive expressions to model can possibly rectify this situation. Alternatively, further studies may result in finding that William's Ostracism model needs to be modified if the feeling attributed to each type of ostracism, and the corresponding survey questions, are not valid.

In the World Comparison Survey 45 of the 49 participants were able to clearly identify which world they preferred. 4 people mentioned that the worlds were identical and thus did not nominate a preference. It is important to note that World 2 was always the Agent world and World 1 was always the Avatar world, but that does not mean World 2 was always experienced second, the world order was randomly assigned. Of the 45 participants, 21 participants preferred World 1 as they thought world 2 contained people that were very hostile and unwelcoming. Another 14 participants preferred World 2 as they thought they were interacting with more realistic people, or were fascinated by the agents changing their emotional state. The remaining 11 participants preferred either World 1 or 2 for miscellaneous reasons such as the music in the room, or the paintings on the walls even though both rooms were physically identical. Here are some of the free text comments of participants.

People in World 1 were much more accommodating, more accepting. People in World 2 changed their emotions more readily, did not seem as hospitable
Participant 13, Inclusion condition

People seemed to be angrier in World 2, or more irritated. I didn't notice any physical differences other than the fact that the expressions of the faces were more aggressive in World 2. I'm not sure but I don't think there was any artwork on the walls in World 1.
Participant 43, Oblivious Ostracism condition

World 2 had guests who displayed expressions on their faces and looked angry all of the time.
Participant 46, Punitive Ostracism condition

In World 2 they were a lot more upfront with their feelings- I could see it on their faces- in World 1 they hid it better.
Participant 32, Defensive Ostracism condition

What is more interesting is that participants went further than identifying the differences. They also identified that the expressions made the guests in the room more hostile, less friendly and mean. This means that the participants were able to interpret the negative emotions being displayed to them. This interpretation led them to conclude that one world was meaner or more hostile than another world. This is largely because of there were more negative than positive expressions displayed by the avatars/agents.

World 2 looked more like real people with expressions
Participant 21, Punitive Ostracism, preferred World 2

World 2 because it is like the real world when you interrupt the people
Participant 2, Defensive Ostracism, preferred World 2

World 1 because even though I was ignored, I didn't feel like they were ignoring me on purpose
Participant 45, Oblivious Ostracism, preferred World 1

World1 - I'd prefer not to see what emotions they are showing
Participant 31, Defensive Ostracism, preferred World1

5. CONCLUSION

It is evident, as Bailenson [1] suggested, that if the agent were perceived to have human-like behavioural qualities, then the human participant would interact with the agent as if it were a human being. Participants were not only able to identify the different emotions depicted on the faces of the avatars and agents but were also affected by them. Most participants were able to detect that the difference between the Agent and the Avatar Cocktail Party world was the constant changing of facial expression in the Agent world and concluded that the Agent world was more hostile than the Avatar world in questionnaires they completed after experiencing both worlds. The hostility that the participants noted was a result of the way the participant interpreted the emotions being depicted, with more negative emotions being clearly identified than positive emotions. These results show that the artificial intelligence component was able to make a significant difference to the experimental condition by creating a more realistic environment to simulate the Punitive Ostracism condition.

Research into emotions and their value is in its infancy! Our research represents a significant contribution to this field by validating previous work conducted in the field by [1,8,9,12] and in doing so, adding to this body of work. The results uncovered by our study establishes the need for further study into modeling emotions in artificially intelligent software agent that interact with a human participant.

In the past emotions have been dismissed as a distraction to the logical, scientific thought process. This is because an individual acting on their emotions was considered to be an irrational being and Artificial Intelligence research is primarily based on the assumption that humans are rational and the simulation of intelligent behaviour should be rational. More recently however, the importance of emotion in human-like intelligence and behaviour has been identified. Moreover it has been identified that different emotional states affect the choice of processing capabilities a human being will employ. If we disregard the role emotions play in the decision making process, we will not be able to develop simulations and planning systems that behave realistically. Without realism their value as learning and training devices will be limited.

In the Agent Cocktail Party World developed and validated in this study, the human participant was able to react to the subtle stimuli received from the depiction of emotions on the agents' faces. While, our emotion based agents were more believable than the scripted avatars, they wouldn't currently pass the Turing test.

However, our experimental results encourage us to believe that removal of the identified current technological limitations of IVR software and hardware, together with development of a more sophisticated emotion based agent architecture will produce a truly realistic and immersive environment that can be applied to many domains.

6. ACKNOWLEDGEMENTS

Thanks to Prof. Kip Williams for his knowledge, guidance and inspiration in this work. Thanks to Dr. Manolya Kavakli and Andy Beall for their support, expertise and input.

7. REFERENCES

- [1] Bailenson, J., Beall, A., Blascovich, J., Raimundo, M. & Weishbush, M. (2000) Intelligent agents who wear your face: User's reactions to the virtual self, Technical Report, Center for the Virtual Environment and Behaviors, Department of Psychology, University of California, Santa Barbara.
- [2] Bates, J. (1994) The Role of Emotions in Believable Agents, *Communications of the ACM, Special Issues on Agents*, July 1994, 94-136.
- [3] Blascovich, J. (2001), Immersive Virtual Environment Technology as a tool for Psychological Science, *Psychological Science Agenda*, November/December issue.
- [4] Canamero, D. (1997), Designing emotions for activity selection, Technical Report, Lego Lab, Department of Computer Science, University of Aarhus Aabogade, Denmark 34 DK-8200.
- [5] Cassell, J. & Vilhjalmsson, H. (1999), Fully embodied conversational avatars: making communicative behavior autonomous, *Autonomous Agents and Multi-Agent Systems*, 2, 45-64
- [6] Ekman, P. (1993), Facial Expression and Emotion, 1992 Award Address, *American Psychologist*, April Vol48, No4, 384-392
- [7] Ekman, P. (1997) Should we call it expression or communication? *Innovations in Social Science Research*, Vol 10, No5 pp 333-344.
- [8] Ekman, P. (1999a) Basic Emotions. *In: Dalglish & Power (Eds) Handbook of Cognition and Emotions*, chapter 3. Wiley & Sons, Sussex, UK.
- [9] Ekman, P. (1999b) Facial Expression. *In: Dalglish & Power (Eds) Handbook of Cognition and Emotions*, chapter 3. Wiley & Sons, Sussex, UK.
- [10] Gruter, M. & Masters, R. (eds). (1986) Ostracism: A social and biological phenomenon. *Ethnology and Sociobiology*, 7, 149 – 395.
- [11] Loomis, J., Blascovich, J. & Beall, A. (1999), Immersive Virtual Environment Technology as a basic research tool in psychology. *Behavior Research Methods, Instruments & Computers*, 31(4) 557-564.
- [12] Oliviera, E. & Sarmiento, L. (2003), *Emotional advantage for adaptability and autonomy*: In: Proceedings of 2nd International joint conference on Autonomous-Agents and Multi-Agent Systems held in Melbourne Australia.
- [13] Rousseau, D. & Hayes-Roth, B. (1997), Interacting with Personality-rich Characters, Technical Report, Knowledge Systems Laboratory, KSL 97-06.
- [14] Williams, K. (2001), *Ostracism- the Power of silence*, Salovey (Ed) Emotions and Social Behavior series, Guilford Press.