BSc AI Thesis - Arne Roland Wijnia

Interpersonal distance in a serious game

Using GSR to measure interpersonal distance reactions

Spring and Summer 2012

Supervisors: Pim Haselager, Franc Grootjen & Inti Brazil

Abstract

Many methods for diagnosing psychiatric patients exist, but few are both entertaining and successful. This thesis tries to contribute to a new diagnosis method that is possibly more motivating than existing treatments: a serious game. Participants have to complete a small task in this game, in which they have to cooperate with an avatar. In two conditions, this avatar is either distant or close to the screen. It is investigated whether or not the avatar's positions make any difference in game performance, implicit GSR reaction, or answers to a questionnaire about the avatar.

Introduction

Cognitive assessments in a forensic context (psychiatric patients convicted to involuntary commitment) use many standard psychological tests (e.g. Stroop task (MacLeod, 1991)) that generally involve very different contexts, stimuli and behaviours than those that called for the assessment in the first place and which, moreover, are far from motivating for the patients. This thesis is part of a larger project that aims to extend the methods for diagnosing psychiatric patients, in a way that is more motivating and equally, if not more, effective. One way to assess psychopaths is looking at brain structures. One interesting brain area is the amygdala, where psychopaths tend to show abnormal activation in functional imaging studies (Adolphs, 2003), (Yang, Raine, Narr, Colletti, & Toga, 2009), which can cause antisocial and aggressive behaviour. Furthermore, abnormalities in the amygdala cause a loss of fear experience and the experience of interpersonal distance (Kennedy, Glher, Tyszka, & Adolphs, 2009). Methods for diagnosing criminal psychopaths should focus on such differences in brain activation. In this thesis a diagnosing tool is investigated that focuses on the lack of interpersonal distance awareness in psychopaths. It is created in such a way that it can also be expanded to focus on social interaction, rather than interpersonal distance. Experiments conducted in this thesis will not have psychopaths as participants, but normal subjects, to provide a baseline.

The new assessment tool is a serious game, a computer game created for a primary purpose other than entertainment. Research in the field of serious games has already been going on for decades (Abt, 1970), and in principle its possibilites are nearly endless. Almost any scenario can be implemented and many daily life situations can be simulated. The amount of scenarios that can be implemented in a time span of half a year, however, is limited.

The possiblities of simulating social situations has already been investigated (Brinkman, Inan, & Mast, 2009), but this thesis will not expand upon that. Instead, the serious game created for this thesis will investigate interpersonal distance reactions, for which an avatar is used. In order to make it a game, there had to be something that involves both the participant and the avatar. A scenario is chosen were the participant and the avatar hold a maze with a ball in it. By tilting the maze in certain directions, the ball moves through the maze. The size of the maze can be easily modified, making it possible to create several conditions, and in each of those the participant is further away from or closer to the avatar. To keep this game challenging, the avatar will not always be cooperating. Disrupting behaviour from the avatar might cause more frustration for the participant. There has not been enough time for this thesis to investigate these frustration patterns, but by including more complex avatar behaviour, a better opportunity for future research in this field has been created. It can potentially be linked to EEG studies, error monitoring and diffusion drift models (Brazil et al., 2009), (White, Ratcliff, Vasey, & McKoon, 2010). Since this thesis is part of a larger project, this can be done at a later stage, with cooperation from assistant professor van Vught (RUG).

The reactions to events occuring in the game can be measured as Galvanic Skin Reponses (GSR). A GSR system measures implicit reactions, that cannot be controlled or manipulated by participants. This makes it an excellent measure for the arousal reactions regarding interpersonal space. The reactions to events regarding the avatar are recorded using a questionnaire. At the end of the game the participants have to fill in a form where they express how they thought the avatar was behaving. This form is an adaptation to a questionnaire by McCroskey and Teven (1999), with questions about the goodwill, competence and trustworthiness of the avatar.

In short, this thesis will look to find an answer to the following questions:

1. Is it possible to implement a working application of a serious game?
2. What effect does distance between the participant and the avatar have on peformance? The amount of games a participant can finish in a certain time is taken as an indication for performance.
3. What effect does distance between the participant and the avatar have on GSR?
4. What effect does behaviour of the avatar have on questionnaire responses?

The main focus is creating a game and investigating whether the distance affects how the participant feels (the GSR signal) and how he performs (the number of games played).

Background

In this section I will go into more detail on Galvanic Skin Response (GSR) signals and the created game.

Galvanic Skin Response

GSR is an old psychological method that is used to measure the skin conductance between two points. Very small changes

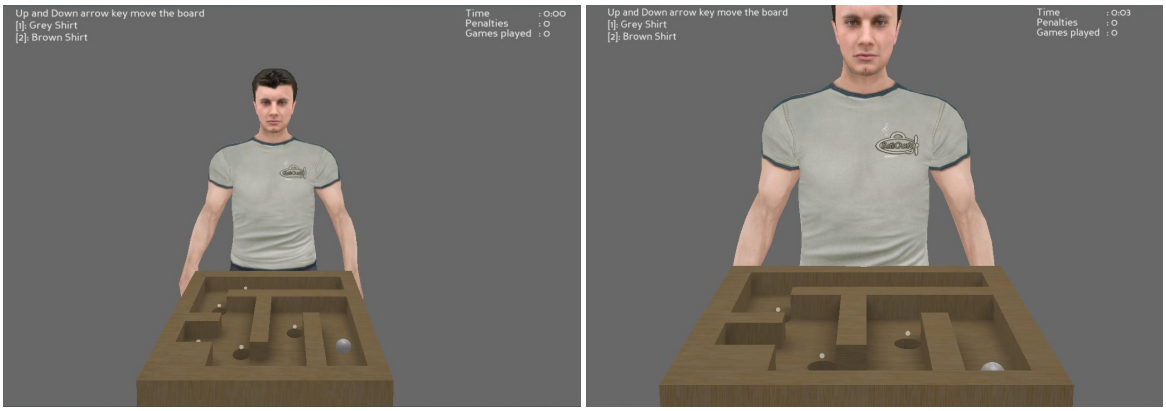

Figure 1: The distant and close condition views

in this conductance, caused by the sweat glands, can be measured. Because the sweat glands are controlled by the sympathic nervous system, one cannot control his GSR reactions. This makes the GSR an excellent tool to measure arousal. Interesting for this research is the fact that psychopaths show less response in GSR studies, so any data collected from this baseline, can be used to compare with the GSR data gathered from psychopaths later on (Tarvainen, Koistinen, Valkonen-Korhonen, Partanen, & Karjalainen, 2001).

Interpersonal distance refers to the notion that humans like a certain distance between each other, when for example in conversation (Hall, 1982). If someone enters your personal space, you feel uneasy and uncomfortable. As noted in the introduction, psychopaths lack this feeling for personal space, making it an interesting diagnosing option. Because reactions to interpersonal space violations are uncontrollable, I expect to find GSR reactions when this happens.

The game

This section describes the game, why certain design choices were made and why other parts are left out. The game is related to interpersonal distance and the question whether it is possible to create a serious game that can be used as a treatment for psychopathy.

Due to limited time and resources a simple game and game environment was chosen. Panda3D was used as the environment. Panda allows for the fairly easy creation of 3D computer games, using Python as its language. Panda offers tutorials which explain all possibilities and one such tutorial proved to be extra useful. This tutorial is about a maze with a small ball in it. By moving the mouse, the maze can be moved and the ball traverses through the maze. This somewhat resembles the scenario that was chosen, making it an excellent starting point for the creation of the game used in this research. To create this game, the original tutorial needed a couple of improvements.

Avatar

To measure reaction to changes in interpersonal distance, a human avatar is added. The avatar holds the board on one side, while the illusion is created that the participant holds the other side. The avatar has a neutral facial expression that does not change throughout the game. Each time the board moves, the avatar has to move accordingly. It is now programmed in such a way that the complete avatar moves with the board, instead of only his hands. This is not optimal and it is currently unclear whether and to what extent this influenced the participants.

Maze

The tutorial already provided a useful maze. The mazes vertical length was doubled, creating a maze that is twice as long. That way, two conditions can be created, one with a long maze, called the distant condition and one with a the original, shorter maze, called the close condition (Figure 1). The maze is small and not difficult to complete in both conditions. The ball can traverse through the maze in anything between 10 and 30 seconds. This allows for many repetitions and thus many occasions to measure responses to certain events.

Zoom

The distance between the participant and the avatar has to change, in order to find a reaction when the avatar becomes too close and invades the personal space. This change is created by adding a zoom each time the game starts. This zoom lasts eight seconds and moves the camera from futher away to the edge of the board, creating the illusion of moving towards the avatar.

Penalties

The game was originally designed for more than only measuring interpersonal distance reactions. It was also created to measure frustration levels and decision making regarding

events in the game. The game holds a few elements that should increase the frustration level of the participant.

The first frustration element is a penalty system. There are several holes in the maze and whenever the ball moves over these holes, the player gains some penalties. He gains one each frame the ball is on these holes. The amount of penalties can increase very fast since frames follow each other rapidly. Furthermore, the holes are placed in such positions, that the ball might get stuck on a hole for a few seconds. The somewhat random nature makes the penalty system a bad measure for game performance and it is not be used in that way, but it is a great way to increase the participants frustration level.

AI and Avatar

To expand this research and include decision making, two avatars and two different AI's are present. At the start of the game, the participant is facing an avatar in a grey shirt, which moves the board according to a cooperative AI. After finishing a couple of rounds, the avatar will move the board according to a less cooperative AI. This second AI slows the game down, by tilting the board very slowly, or by moving the ball to a wrong direction. Furthermore, the ball is moved over the holes, thus giving the participant more penalties and increasing his frustration. At a certain point, only the second, inefficient AI is used. This should push the participant to make the desicion of changing the avatar. The second avatar has a brown shirt and a different face, but the AI pattern is the same. So after a few games, this avatar will appear to be disrupting the participant as well. The exact pattern is as follows: the first two rounds are with a cooperative AI, then the disruptive, then a round cooperative and each round following will be with the disruptive AI. This pattern is resetted each time the avatar is switched. To prevent participants from continuously changing the avatar, they lose a 'games played' each time they switch. This also damages your performance, since the number of 'games played' is the used performance measure. This element will make the decision to swith the avatar more difficult: not switching means facing a disruptive AI, but switching means losing a 'games played'.

EEG

The penalty system and the different AI's and avatars belong to a non performed part of the experiment. Orginially, reactions to these elements were to be measured but due to time and technical limitations, no EEG signals could be recorded. This made it also impossible to investigate the frustration and decision patterns, making the frustration elements superfluous. These elements are still in the game however, making it possible to perform the EEG experiments in later research. The major technical problem was the placement of markers in the GSR data.

Markers

One can only succesfully analyse GSR or EEG data, if it is exactly clear at what point an event occured. A stream of GSR data is worthless without markers pointing out the places were certain ingame events took place. The markers had to be programmed into the game and transported to the Biopac GSR system by use of the parallelport. Due to some technical issues it proved impossible to place these markers. However, the system allowed for the placement of manual markers. Each time an event occurs, a key combination has to be pushed and a marker is placed. This is less accurate than automatic markers but GSR signals are relatively slow which made it possible. When recording faster signals, like for example EEG, then this would not have been an option. The markers are placed at start and stop of the game, each time a zoom begins, a zoom ends and as the player switches his avatar.

Experiment

This section describes the experiment. With this experiment, it is tested if distance between the participant and the avatar matters in performance. GSR signals are also measured and the participants have to fill in questionnaires.

Participants

A total of 9 healthy male students were tested on the experiment. All of the participants were of age 18-22. Two of these were left-handed and all of them had correct eye sight.

Apparatus

A standard colour monitor is used to display the game to the participant. A standard keyboard is used, connected to the monitor displaying the game. The up and down arrow key are used to move the board, with the invert-Y technique. This means that the down arrow will push the front of the board up and the up arrow will push the back of the board up. The '1' and '2' number keys are used to switch avatars.

Recording

The data is collected with the Biopac system. Two GSR transducers are placed on two of the participants fingers, the index and middle finger of the left hand. The cavities of the GSR are filled with electrode gel to obtain accurate readings. The GSR is used to measure differences in the skin of the participants, caused by, for example, perspiration..

Procedure

During the experiment, the participant is sitting in front of the computer. He has to finish two different conditions. At the

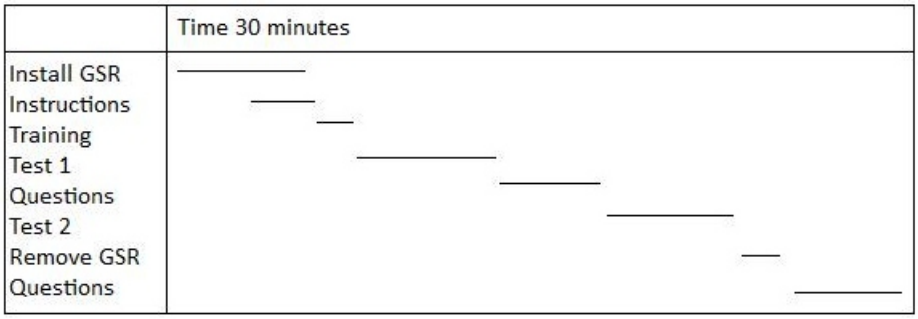

Figure 2: The timeline of the experiment

start of the experiment the GSR is attached to his fingers. He then recieves some general instruction as to what he needs to do (Appendix B). The participant finishes the two conditions, each directly followed by a questionnaire. The total time for the experiment is roughly 30 minutes (figure 2).

The experiment has several phases: preparation, training, testing and filling in the questionnaire.

Preparation

During the preparation, the GSR is positioned on the fingers of the participant. He is seated in front of the computer running the game.

Before the GSR experiment started, each subject did two tests. The first test was labeled the blow test. The participant had to blow on his left hand, were the GSR was attached. If everything was working correctly, the GSR would show a clear spike after the blow. This was found to be the case for every participant. The second GSR test was labeled the cough test. The participant had to cough and if everything was working, the GSR signal would show a clear spike and again this was found. After both tests were succesfully completed, the participant would get a small training.

Training

During the training, the participant receives instructions about the game and the two conditions (Appendix B). The goal of the game is explained. To make sure that the participant understands what he is doing, we start up a first test round. During the test round, the participant gets to see the game and use the controls for a few seconds. That way, the participant understands the controls and the mechanics.

Test

During the actual test, the participant has to play the game several times. Both the first and the second condition will be used. From the nine participants, 5 started with the first condition (long maze) and 4 with the second condition (short maze). After each time they get the ball through the entire maze, the camera will zoom again. This camera zoom has to occur at least 10 times. After you switch the avatar, the camera will also zoom. Each participant plays for about 5.30 minutes or more, given that the camera has to zoom at least 10 times.

Questionnaire

After the first test, the participant has to answer a questionnaire. Then the game will be played a second time, after which the GSR is removed and the participant has to fill in a form again. The questionnaire has to be filled in a total of four times for each participant, because there are two different conditions and two different avatars. Appendix A holds the questionnaire.

Results

The results gathered are from both conditions. The condition with the larger maze, where the avatar is distant, is referred to as condition 1 and the condition with the the shorter maze, where the avatar is close to the screen is referred to as condition 2.

GSR Results

Each participant experienced at least 10 zooms in each condition. A zoom followed after an avatar switch, or after a game was played. A typical GSR reaction had to occur during the zoom to classify it as a reaction to the changing interpersonal distance. Such a reaction is shown in figure 3. In a lot of the cases however, no reaction could be measured at all. A lack of typical reaction is shown in figure 4. Anything in between can occur as well. It is sometimes hard to classify whether or not a signal is a reaction to the zoom, or still a reaction to an event occuring prior to the zoom, or an event occuring right after the zoom. Multiple smaller spikes during the zoom also occured, which are also difficult to interpret. If it resembled mostly the spike seen in figure 3, it is classified as a reaction, but very uneven signals are not classified as a reaction.

Other methods for analysing the data should also be considered, but due to limited time, these methods are not used

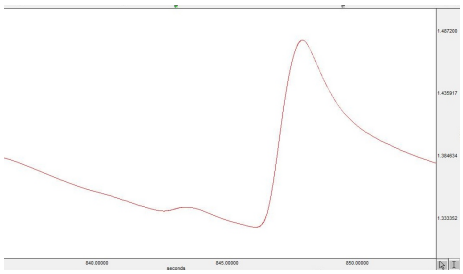

Figure 3: A typical GSR signal as response to a stimulus. The two inverted triangles mark the start and end of the zoom.

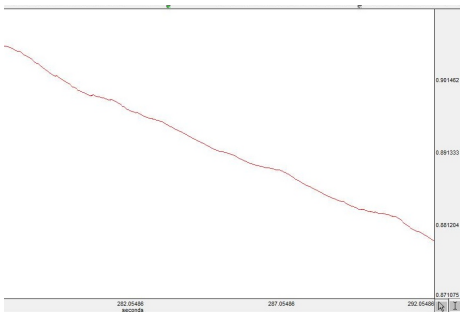

Figure 4: The two inverted triangles mark a zoom, but no typical reaction can be seen.

in this thesis. Determining the onset of the signal can also be done with more mathematical methods. One way to do so is by establishing a threshold, which is for example half of the average peak amplitude for each participant. Every spike above this threshold can than be classified as a reaction. The results shown here, are all gathered only by visual inspection.

	large maze	short maze
	% responses	% responses
P1	30.76	33.33
P2	16.67	28.57
P3	30	30.76
P4	38.46	62.5
P5	50	33.33
P6	18.18	33.33
P7	25	16.67
P8	75	66.67
P9	66.67	46.67
Total	41.23	39.67

Table 1: The percentage of zooms that triggered a response

During condition 1 (with the larger maze) a total of 114 zooms occured, distributed over the nine participants, with 10 zooms being the minimum and 16 the maximum. Of the 114 recorded zoom reactions, 67 showed no response to the zoom, 47 did. During condition 2 (with the shorter maze), similar results were found. 121 zoom reactions were recorded in the second condition. Of these zoom reactions, 73 had no GSR response and 48 had. Even though the maze has a different size in the two conditions and the distance between the participant and the avatar differed, the GSR reactions are nearly the same in both conditions. This means that being closer to the avatar has in general no effect on the GSR signal over all subjects.

Because not every participant played the same amount of games, and the amount of games even differed between the conditions, it is also useful to look at the relative amount of responses. These can be found in table 1. The absolute results can be found in table 2. Participants have different scores, both in the absolute and in the relative table. I have devided them into three groups that are based on specific response patterns, retrieved from the absolute results.

1. P1,P5,P7. These participants showed more or the same amount of responses in condition 1, compared to condition 2. This makes this group the most suprising.

2. P2,P3,P4,P6. This group shows more responses in condition 2 compared to condition 1. For P2, P4 and P6, the amount of responses even doubles. Moreover, P4 is extra interesting since the amount of responses exceeds the amount of 'no reponses' in condition 2.

3. P8,P9. These participants have more responses than 'no responses' over all. P8 and P9 show more respones in condition 1. P8 shows the same effect in condition 2, P9 does not.

Group 1 is behaving exactly opposite of what was expected. These three participants do not react more in condition 2 than in condition 1, while the avatar in condition 1 is futher away. These differences are not that big though. P1 does not have any difference at all, and P5 and P7 only have a difference of 1 reaction. Relatively seen, the difference is however bigger for P5 and P7. They have both more 'no responses' in condition 2, meaning that they reacted less at the changing distance, even though they saw more or just as many zooms in the second condition.

Group 2 consists of 4 participants that show more responses in condition 2. However they are not that similar. P3 is relatively seen almost a member of group 1. He does show one more response in condition 2, but also two 'no responses'. So in condition 2 he only has 0.76% more reactions. P2 and P6 are the prototype members of this group: more responses in condition 2 and just as many or less 'no responses'. Both P2 and P6 double their responses from 2 to 4. At last P4 is the one participant behaving as expected. His number of responses increases in condition 2, while the number of 'no responses' decreases. He doubles his reponses both absolute (5 opposite 10) and (almost) relative (38% opposite 62%). P4 could also be classified as a member of group 3: he has overall more responses than 'no responses'.

Group 3 is a group with participants that show more responses than 'no responses'. P8 and P9 are members of this group and P4 could be placed here as well. Both P8 and P9

	large maze		short maze	
	responses	no responses	responses	no responses
P1	4	9	4	8
P2	2	10	4	10
P3	3	7	4	9
P4	5	8	10	6
P5	6	6	5	10
P6	2	9	4	8
P7	3	9	2	10
P8	12	4	8	4
P9	10	5	7	8
Total	47	67	48	73

Table 2: The GSR results for each participant

however, show less responses in condition 2. This could also make them members of group 1, but the large amount of responses in both conditions indicate that P8 and P9 had much more GSR reaction in general.

Another interesting thing can be seen when looking at the relative results. Five of the nine participants actually score a higher percentage in the second condition, implying that most participant did react more in condition 2 to the changing interpersonal distance. However, this difference is very small for P1 and P3 (smaller than 5%) and overall the difference between the two conditions only differs 2% in favour of condition 1.

The relative results are also compared using a t-test. With a F value of 0.001 and a p-value of 0.981, it is very clear that there is no significant difference between the two conditions.

Performance results

The performance is measured by the amount of games each participant finished.

	large maze	avatar switches	short maze	avatar switches
P1	11	2	9	1
P2	8	2	8	3
P3	9	1	8	3
P4	10	1	7	4
P5	11	0	9	3
P6	11	0	8	2
P7	10	1	8	2
P8	11	2	9	1
P9	8	3	8	3
Total	89	12	74	22

Table 3: The games played by each participant and the total

In condition 1, 89 games were finished by the participants. In condition 2, only 74 games were finished, a difference of 15 games. The table below shows the results for each participant. Participant 2 and 9 had the same amount of games played in both conditions, all the other participants finished more games in condition 1 than in condition 2. The biggest difference is three games. One reason for the smaller amount of games played in condition 2 is the fact that the participants switched the avatars more in condition 2, almost twice as much (12 and 22).

Note that some players started with condition 1 and some with condition 2, but that did not cause a difference for the amount of avatar switches. Only P1 and P8 switched more in condition 1, and P1 started with that condition, where P8 did not.

Questionnaire results

Each participant filled in four questionnaires, this questionnaire can be found in Appendix A. For each question the participant fills in a number between 1 and 7, where 1 means that he does not agree at all with the statement and 7 means that he agrees a lot. Question number 2,3 and 7 are reversed questions, if a participant has given a high mark here, he is actually negative about the avatar. A high mark with the other questions means that he is positive about the avatar. If a participant does not switch he cannot grade the second avatar. This is the case for participant five and six, who both did not switch in the first condition.

The questions belong to one of the three dimensions of McCrowsky. 'Cares about me', 'Has my interests at heart', 'Self-centered' and 'concerned with me' belong to the dimension of goodwill. 'Untrained', 'Inexpert', 'Incompetent' belong to the dimension of competence. At last, 'Untrustworthy' and 'Phony' belong to the trustworthiness dimension. In general, no remarkable answers can be seen. Most participants give neutral grades (a 4) or low grades (3 and 2). No apperent difference between the two conditions can be seen and three t-tests confirm this. For these tests, the average was taken from both shirts in both conditions. If the participant did not switch, the data was taken from the avatar that was used instead of the average. Then for each participant, the answers to questions regarding the same dimension were added to get one single number for each participant in each dimension.

The first dimension, goodwill, is the most difficult one. It contains the reversed questions 2,3 and 7 and question 6. The answers to questions 2,3 and 7 first need to be reversed again,

Participant	P2	P3	P4	P6
Untrained	4.5	4	4.5	6
Cares about me	4.5	5	3	6
Has my interests at heart	5.5	3.5	3	5
Untrustworthy	4.5	4	3.5	4
Inexpert	4	5	4	4
Self-centered	4.5	3.5	4.5	5
Concerned with me	5.5	4.5	2.5	3
Incompetent	3.5	4	4	5
Phony	4	5.5	3	5

Table 4: The average of condition 1 for P2, P3, P4 and P6

Participant	P2	P3	P4	P6
Untrained	4.5	3	4.5	4
Cares about me	4.5	5	4	4.5
Has my interests at heart	4.5	3.5	4.5	3.5
Untrustworthy	4.5	3	3.5	2.5
Inexpert	4.5	3.5	4.5	3
Self-centered	4	4	4	3.5
Concerned with me	4	5	3	4.5
Incompetent	5	3	3.5	3
Phony	4	5	1.5	3

Table 5: The average of condition 2 for P2, P3, P4 and P6

to make them comparable with question 6. Each answer was withdrawn from eight, as to reverse them again. All the answers were added up for each participant and a paired t-test was performed. The df was 8 and the sig. (2-tailed) was 0,584 showing no significant difference between the two.

The second t-test used competence dimension answers. Once again, all the answers were added up to get one number for each participant. The results show the same as in the goodwill test: no significant difference can be found (sig. (2-tailed): 0,473 and df: 8)

The third test used trustworthiness answers. The results show the same as in the goodwill and competence test: no significant difference can be found (sig. (2-tailed): 0,276 and df: 8) These tests show that there is no significant difference between the answers of either of the dimensions regarding the two conditions.

Given the GSR and performance results, it is interesting to look at individual participants, especially P2, P3, P4 and P6, because these participants showed more responses in condition 2 compared to condition 1.

Tables 4 and 5 show the average over the two shirts for both conditions for P2, P3, P4 and P6.

The averages over both avatars (note that P6 did not switch in the first condition, hence his average is the same as his score for the grey shirt) do not show many extremes. The scores for P2 are almost all neutral, either 4.5 or 4, for both conditions. P3 seems to be a bit more negative in condition 2, rating almost every question 1 point more negative. P4 differed most in reactions between the conditions regarding the GSR results, but no big differences between the two conditions can be seen in the questionnaires. Where he is more negative on some questions, he is more positive on others. At last, P6 grades almost every question higher in condition 1, which is puzzling, because question 2 and 3 should be graded lower (like he did for question 7) to be consistent with a more positive attitude. He is more positive in condition 1 about every aspect of the avatar, except for 'cares about me' and 'has my interests at heart'. On the other two questions of the goodwill dimension, 'self-centered' and 'concerned with me', he is 1.5 point more positive.

Conclusion

An explanation for the results found is given in this section. Starting with the GSR, performance and questionaire I will end with the conclusion to the main question: is it possible to implement a working serious game?

Explanation for GSR results

The result of the GSR data shows that there is little response to changes in interpersonal distance. However, P2, P3, P4 and P6 showed more response in condition 2, where the avatar was closer. On the other hand, only P4 and P8 showed more responses than 'no responses' in condition 2. When looking at the relative amount of response, P1 also has more response in condition 2. This difference is not significant however. The conclusion is thus that there was no significant difference in the amount of responses between condition 1 and condition 2. The large amount of 'no responses' in both conditions indicates that there was little response to changes in interpersonal distance in this game. There are several possible explanations for this. I list them here.

1. The first one is that participants do not experience a personal space with a virtual person. So whether or not an avatar is appearing really close or really for away does not matter, since you do not experience him as being either in your personal space or not.

2. A second explanation is that the avatar was too far away. It could be that he was never close enough to give the participant the feeling that his personal space was invaded. So even the small maze could still be too large.

3. Another explanation is that the avatar behaved not human enough. Perhaps the fact that the entire avatar moved with the board, instead of just his hands, made it too obivous that it was just a computer generated being. Furthermore, the avatar did not respond to anything happening in the game, where a real person would do so. The lack of facial expression and verbal interaction could have made the avatar too unreal.

4. Perhaps the task was too distracting. The game that had to played proved difficult and frustrating for most participants. They tried their very best to perform properly, but failed mostly because of the bad AI condition. The only moment of rest they had, was during the eight second zoom. So it could be that they did not pay attention at all to the avatar and whether he was close or far a way. Several participants even reported that they did not notice any change between condition 1 and condition 2, or any change between the grey shirted avatar and the brown shirted avatar. This could mean that they were so focused on the game, that they never realised that there even was an avatar.

Explanation for performance results

The performance did show some difference between condition 1 and condition 2. In total, 89 games were finished in condition 1 and 74 in condition 2. This difference can partially be explained by the fact that the participants switched their avatar a lot more in condition 2. An explanation for both the fact that the amount of games played and the amount of avatar switches differs between the two conditions, is the fact that the AI differed between the two conditions. Each condition had a cooperative AI which was roughly the same. The more disrupting AI however, differed between the two conditions. Due to time issues, this difference could not be resolved. During tests, one aspect of the disrupting AI in the short condition proved to be a large challenge for the participants. The avatar would at one point, if the participant did not react fast enough, steer the ball to a corner of the maze. The only way the AI would then futher react was if the participant moved the ball all the way back and over a hole. This cost a lot of time and most participants did not want to do this and instead switched their avatar.

It is this and perhaps other differences between the AI's that could have cost the participants more time and more avatar switches to finish the short maze. It seems unlikely that the avatar had any effect on performance. The avatar did not effect the GSR signals and the AI is more likely to play a role in game performance.

Explanation for questionnaire results

The behaviour of the avatar did not matter in questionnaire responses. Most participants gave neutral or negative marks for the questions. When asked to fill in the questionnaire, most seemed not certain what answers to give. When asked why they had trouble grading the avatar they reported that they did not see the avatar as someone with character traits like 'trustworthy', or 'inexpert'. This can mean that the questionnaire was wrong and different questions should have been asked. More likely is the fact, and this seems to be supported by the GSR and performance results, that they did not see the avatar as human, but simply as an element of the game that was, contrary to the maze and ball, unimportant. Thus they did not pay attention to him and could not grade him. Because the questionnaire was a part of the experiment, participants had to grade the avatar. An explanation for the negative grades can be the fact that they were still frustrated about the game and blamed the avatar for the bad results they had gotten. The somewhat similar results from the questionnaire in both conditions, indicates then that the participants were equally frustrated about how the game went and that not the distance of the avatar mattered, but only its behaviour.

P2, P3, P4 and P6 showed more GSR response to the changing interpersonal distance in condition 2. However, these participants did not react in a same manner regarding the questionnaire. For example, P2 was neutral in both cases, while P6 was more negative about the avatars in condition 2. Suprising is also the differences within a dimension. For example, P6 graded the questions in the goodwill section very inconsistent, while these questions are only slightly different. These results further confirm the theory that the participants did not spend enough attention to the avatar to grade him properly.

It seems that the interpersonal distance did not affect the way the participants felt about the avatars. More likely is that the frustration generated during the game influenced the answers, but new research will have to investigate and confirm this.

Serious game

The results show that a succesful serious game that can measure reactions to interpersonal distance is not implemented. The question remains if this is because the game itself was flawed, or if it is impossible to measure interpersonal distance reactions in games at all, regardless of the game. Future research will have to determine if this is the case.

Future research

This thesis is pilot work for a larger project. This means that a lot of work can still be done in the future, aided by the results gathered here. This research can be expanded and improved, or start another similar research.

Improving the game

The results show that little reaction to changes in interpersonal distance were measured. To make it possible to see these reactions, the game can be changed in several ways. Overall, it seems a good idea to involve the avatar more in the game. This means that the game should revolve more around the avatar, rather than the task the participant has to complete.

The avatar could be given a greater role in the game. This includes some improvements in movement (not the entire body, but just its arms). Facial expressions can be added and some verbal interaction. At the very start of the game,

a small introduction round could be added, were the avatar tells his name and maybe explains the purpose of the game. The interaction with the avatar will likely increase the particpants awareness of the avatar and perhaps make him more suspectible to changing interpersonal distance.

The game can also be simplified. The penalty system and avatar switch was added to measure frustration increase and decision making, but for the sake of just interpersonal distance research, these elements can be left out. The disrupting AI is also not a neccesary element, leaving only a cooperating AI and a simple maze.

Different games

Other game scenarios were considered before this one was chosen. With more time and rescources, these could be implemented.

Being the ball is considered as an option of a game similar to this one. To have human interaction, the maze could be replaced by groups of humans. That way, you get the idea of manouvering through a party, or crowded room. Each time you bump into a group of humans, or come too close to them, they should provide some interaction.

More complex games, that involve other elements than just interpersonal distance and social interaction are also diagnosis possibilities. These can include people annoyingly staring at the participant, complex decision making and anger evoking situations. New research will have to determine whether or not it is possible to get useful reactions from such games.

Appendix A - Questionnaire

Thoughts about my opponent:

1. Untrained 1 2 3 4 5 6 7 Trained
2. Cares about me 1 2 3 4 5 6 7 Doesn't care about me
3. Has my interests at heart 1 2 3 4 5 6 7 Doesn't have my interests at heart
4. Untrustworthy 1 2 3 4 5 6 7 Trustworthy
5. Inexpert 1 2 3 4 5 6 7 Expert
6. Self-centered 1 2 3 4 5 6 7 Not self-centered
7. Concerned with me 1 2 3 4 5 6 7 Not concerned with me
8. Incompetent 1 2 3 4 5 6 7 Competent
9. Phony 1 2 3 4 5 6 7 Genuine

Appendix B - Instructions

You are about to play a computer game, called ball-in-maze. It is your goal to move the ball through the maze, by tilting the board in the vertical direction. The opposing guy will move the board horizontally, so a certain amount of cooperation is needed to succeed. As soon as the ball has traversed through the entire maze, it is teleported back to the beginning and you gain a 'games played'. Each time you finish a round, the camera will zoom in again on the scene. Certain areas of the maze should be avoided, these are the holes. Each time the ball moves over such a hole, you gain penalties, according to the time you stay on the hole. If you decide that your current coworker is no longer performing properly, you can switch to a guy in a brown shirt. You can switch whenever you want, but not when the camera is zooming in. You can also switch back to the guy in the grey shirt. Each time you switch, you lose a 'games played'.

References

Abt, C. (1970). *Serious games*. New York: The Viking Press.

Adolphs, R. (2003, March). Cognitive neuroscience of human social behaviour. *Nature Reviews — Neuroscience*, *4*, 165 - 178.

Brazil, I. A., Bruijn, E. R. A. de, Bulten, B. H., Borries, A. K. L. von, Lankveld, J. J. D. M. van, Buitelaar, J. K., et al. (2009). Early and late components of error monitoring in violent offenders with psychopathy. *Biological Psychiatry*, *65*, 137-143.

Brinkman, W.-P., Inan, F., & Mast, C. A. van der. (2009, April). A virtual environment to create social situations: First step to a virtual reality exposure therapy system for social phobia. *Euromedia Conference*, 103-107.

Hall, E. T. (1982). *The hidden dimension*. Anchor Books.

Kennedy, D. P., Glher, J., Tyszka, J. M., & Adolphs, R. (2009, October). Personal space regulation by the human amygdala. *Nat Neuroscience*, *12*, 1226-1227.

MacLeod, C. M. (1991). Half a century of research on the stroop effect: An integrative review. *Psychological Bulletin*, *109*(2), 162-203.

McCroskey, J. C., & Teven, J. J. (1999). Goodwill: A reexamination of the contruct and its measurement. *Communication Monographs*, *66*, 90-103.

Tarvainen, M. P., Koistinen, A. S., Valkonen-Korhonen, M., Partanen, J., & Karjalainen, P. A. (2001, October). Analysis of galvanic skin responses with principal components and clustering techniques. *IEEE Trans Biomed eng*, *48*, 1071-1079.

White, C. N., Ratcliff, R., Vasey, M. W., & McKoon, G. (2010). Using diffusion models to understand clinical disorders. *Journal of Mathematical Psychology*, *54*, 39-52.

Yang, Y., Raine, A., Narr, K. L., Colletti, P., & Toga, A. W. (2009). Localization of deformations within the amygdala in individuals with psychopathy. *Arch Gen Psychiatry*, *66*, 986-994.