Deception Detection by Humans in the Iterated Prisoners Dilemma

Wietse Smid
University of Twente
P.O. Box 217, 7500AE Enschede
The Netherlands
w.b.t.smid@student.utwente.nl

ABSTRACT
For humans, lying is a complex thing to detect, both when lying and when being lied to. While telling a stone cold lie might be difficult for humans, it is relatively easy for a computer to do so. People pick up on lying peers by some involuntary telltale signs and emotions that are easy to mask on a computer image as it does not suffer from such human traits. Thus when playing a game, actions can speak louder than words; people might tell one thing, but do another.

This study examines whether humans can pick up on an Embodied Conversational Agent (ECA) portraying emotions belonging to a certain strategy, when actually playing the opposite strategy and if they act accordingly. This was done by having participants play a game of the Iterated Prisoners Dilemma (IPD) against an ECA opponent. The strategy of the ECA and the fact whether the ECA displayed conflicting emotions or emotions that are consistent with the played strategy were varied between participants. Afterwards participants were asked what they thought of the ECA they had just played against through a short questionnaire.

An online game was developed and 42 persons participated. Out of those 42 participants, 39 completed the questionnaire at the end. The other three participants ended the experiment after having played the initial game.

The results showed no significant differences between the researched groups, both in terms of emotions (the action taken) and in terms of cognition (the questionnaire).

Some filtration and interpretation of the data, combined with results from previous research into the subject, might warrant further research, as the subject itself is still very much topical in today’s science.

Keywords
Deception Detection, Human cooperation, Embodied Conversational Agents, Iterated Prisoners Dilemma.

1. INTRODUCTION
Humans have a long history of trying to recreate themselves. With the advent of the computer in the 20th century, this endeavor has turned more and more into a lifelike reality. We have been able to create continuously better imitations of ourselves, and nowadays much of the research this field is focused around Embodied Conversational Agents (ECAs).

ECAs are an effective way for humans to interact with computers. It is shown that in interaction the latter are treated by humans as if they were another human, and that humans react to them in ways they would react to other humans [2]. ECAs also offer a variety of ways to simulate nonverbal human-like behaviors such as facial expressions and emotions.

Research has shown that when humans are made to play a game with ECAs, they react to the emotions and expressions shown by the ECA. This is shown in particular with the Iterated Prisoners Dilemma (IPD) [5], which will be discussed below.

1.1 Iterated Prisoners Dilemma
The Prisoner’s Dilemma (PD) is a non-zero sum game in which two players have to make a decision whether to cooperate with the other player or not, without conferring with that player. The payoff table is typically something like Table 1.

Table 1. Typical payoff table for the Prisoners Dilemma

<table>
<thead>
<tr>
<th></th>
<th>Player 1</th>
<th>Player 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cooperate</td>
<td>P1:5   P2:5</td>
<td>P1:2   P2:7</td>
</tr>
<tr>
<td>Defect</td>
<td>P1:7   P2:2</td>
<td>P1:4   P2:4</td>
</tr>
</tbody>
</table>

Even though the best outcome for the players would result if both players cooperate where each gets 5 points, the expected value for one single player is the best if that player defects (5.5 points) compared to when the player cooperates with his opponent (3.5 points).

In the IPD, the PD experiment is repeated several times, with the players seeing the results after each round. This opens up the game to cooperative strategies, where players can achieve synergy and reach a higher combined payoff through continuous cooperation.

1.2 Related Work
The Iterated Prisoner’s Dilemma has been a fruitful topic for research into how humans behave when faced with decisions without having access to all the possible information. Many
studies have been done into finding perfect strategies for the IPD. At first focusing on the mathematical aspects, and later also including emotions and simulated humans.

Around 1980, Axelrod organized a few tournaments inviting people to invent their own strategies to see which would fare best in a competition [1]. He found that many of the devised strategies worked fine, but that one stood out: Retaliator, or tit-for-tat. This strategy begins with cooperation on the first move, after which it copies the last move of the opponent, thus utilizing the principle of reciprocity.

Later, more research was done in finding new strategies for the IPD, including the work of Nowak et. al. [10], who found a strategy that could reliably beat tit-for-tat that they called: Pavlov. Pavlov is based on the principle of win-stay lose-shift. Similar to tit-for-tat it uses a simple rule set to define its behavior. Pavlov only cooperates if both players chose the same action in the last round. The result of this is the ability to exploit unconditional cooperators and the fact that two Pavlov agents will never fall into a circle of alternating defection, something that tit-for-tat falls prey to.

Darwen and Yao then evolved new strategies with genetic algorithms [3]. This showed that cooperation becomes the more dominant strategy over time, but notes that unconditional cooperation can easily be taken advantage of, and that for the generated strategies to be strong and robust (capable of doing well versus any other strategy), the initial seeding should have known strategies of the sort.

Further research started focusing on the non-mathematical facets of the IPD. In their work, de Melo et. al.[5] demonstrated that when humans are made to play a variant of the IPD against a computer opponent in the form of an ECA that showed emotions similar to humans, they tend to act differently than when playing against a regular computer opponent. They furthermore showed that the type of emotion of the ECA also has an effect on the actions taken by the human player.

This research extended the work of de Melo et. al. [6] to include lying computer opponents, probing whether humans can be fooled by a lying computer. Lying is simulated by showing emotions that one would not expect if told the underlying strategy of the opponent. A competitive opponent for example, would normally be happy with defecting to a cooperating human, but a lying competitive opponent might not be, and might even show anger at the arisen situation.

2. PROBLEM STATEMENT

When playing the IPD with an ECA, the decisions a person makes are affected by the emotions shown by the ECA after previous outcomes, in part because of the character that is attributed to the ECA by the subject [5].

However, this has only been shown in situations where portrayed emotions and the strategy of the ECA are consistent with each other. This means that when an ECA shows joy after betraying the subject, its decision making algorithm is also aimed towards that outcome [4]. This positive correlation between emotion and action might contribute to the correct attribution of a character to an ECA by subjects.

The emotional part of the ECA is, at least partly, used to infer a character. This inferred character is then linked to an underlying strategy under the assumption that the portrayed emotions are consistent with the strategy. This follows from the fact that the decisions that are made are influenced by this emotional part [5]. But it poses the question whether humans actually use the choices made by an ECA in their inference of the character and strategy, and whether they can tell that an ECA has conflicting emotions or not.

Rehm and André [11] suggest that this inference of emotional character is done subconsciously. Therefore it is interesting to see if the emotions of the players, represented by the actions they take, deviate from their cognition, represented by the answers that they give in the post-game questionnaire.

Following de Melo et. al. [5], human cooperation rates are used as an effective way to infer the attitude of a person towards an ECA. Humans logically tend to cooperate more with a cooperative person, as the payoff is better. The risk for this high payoff, however, is too big when a person believes he is playing against a competitive ECA i.e. one that defects relatively often, as this would lead to big losses.

2.1 Research Questions

Following the problem statement, this research will focus on answering the following questions:

1. Can humans correctly infer an agent’s strategy when its actions and emotions are conflicting?

To answer this question several sub questions have to be answered that support it:

Humans might only subconsciously react to the emotions of an ECA. If they are asked afterwards for their thoughts, these might be in conflict with what they showed from their actions.

1.1. Do cooperation rates of humans in the IPD differ between ECAs with emotions consistent with their strategy and ECAs with conflicting emotions?

This question looks at the mere reaction part of the humans. Do they really choose different actions when playing against the different types of ECAs.

1.2. Do humans actively recognize ECAs with conflicting emotions?

Some of the factors that are involved in making the decision might only show up when an ECA is clearly cooperative or competitive. It is important to know whether the effects are universal, or if they are limited to a specific type of opponent.

1.3. Are the answers to questions 1.1 and 1.2 true for both cooperative and competitive ECAs?

3. METHOD

To investigate whether people can correctly infer an agent’s strategy when its actions and emotions are conflicting, a web application was built in which participants played an IPD against one out of a pool of pre-programmed ECAs.

In a between subjects design, participants were put up against a computer player which had one out of three strategies and one out of two emotion types. This resulted in a 3x2 between subjects experiment, allowing for comparison of the different strategies and emotion types.

For the research a slightly different version of the IPD was used that was also used in the study by de Melo et. al. [6] called the Investment Game (IG). This version does not present the participant an isolated choice between cooperating and defecting but it is situated in a business setting. Participants are told that they have to invest some sum of money into either Project Blue or Project Green and that the ECA will also invest in either of these projects, depending on both investments, some reward is given to the player (See Table 2).
Kiesler et al. [8] earlier used this type of IPD as an effective way to eliminate any prior unwanted feelings of being “against” a person. It also removes the somewhat negative notion of negotiating punishment. Participants are also told that their goal is to gain the highest collaborative profit, but that they should at the same time try to beat their opponent. Furthermore, they are informed that the ECA has a character itself.

They then played 25 rounds of this investment game against the ECA, all the while seeing their own and the ECAs last action, and the total running score. Figure 1 shows the standard GUI for this part of the experiment. On the left hand side is the ECA, here seen as showing a neutral emotion. These emotion movies where made using the “Responsive Face Demo” from the New York University media research lab [9].

On the top right is the payoff table in textual form, the middle right shows the current status of the game: last round result and running totals. Finally on the bottom right are the choice buttons for the game.

Afterwards, participants are presented a small questionnaire, with questions designed to see what they thought about the ECA and its emotions and strategy.

3.1 Strategy
Following de Melo et al. [6] an IPD setup was used where the first 5 actions of any of the agents are a fixed sequence: cooperation, cooperation, defection, defection, cooperation, which translates to projects Green, Green, Blue, Blue and Green in this setting. This establishes a baseline where the participant has an opportunity to get familiar with the interface and where they can see a few of the reactions of the ECA and get familiar with them; this is however not made explicitly clear to the participants as this would lead to them being able to play perfectly for those first few moves.

After the first 5 rounds, 20 rounds of programmed behavior specific to that ECA’s character will take place. There will be 3 possible strategies that an ECA can act according to: Cooperative, Competitive and Neutral. Table 3 shows how many of the 20 possible choices for the game will be cooperative. These numbers are fixed, this is done to prevent any chance incidents where a cooperative strategy acts competitive and vice versa.

Table 3. Cooperation Rates

<table>
<thead>
<tr>
<th>Strategy</th>
<th>Cooperation Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cooperative</td>
<td>15/20</td>
</tr>
<tr>
<td>Neutral</td>
<td>10/20</td>
</tr>
<tr>
<td>Competitive</td>
<td>5/20</td>
</tr>
</tbody>
</table>

These are not ideal strategies like Tit-For-Tat or variations to that theme [10]. This is explicitly done as the research questions focus on reactions to emotions, and this way there is certainty to the cooperativeness of an ECA, even though the strategies might seem random and/or suboptimal.

3.2 Emotion
Next to the strategies, the ECA will have an emotion type. Here there are two distinct possibilities. The ECA can have emotions consistent with its strategy, or it can have conflicting emotions. Table 4 shows the reactions of these combined personality types to certain outcomes of a single dilemma. The responses are again taken from de Melo et al. [4]. The reaction for the neutral strategy type is always neutral, and the reaction for the conflicting emotion type is the inverse of that of the consistent emotion type. For example, Cooperative-Conflicting behaves the same as Competitive-Consistent.

Table 4. Emotional Response for consistent emotion types

<table>
<thead>
<tr>
<th>Strategy</th>
<th>Outcome</th>
<th>Response</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cooperative</td>
<td>Both cooperate</td>
<td>Joy</td>
</tr>
<tr>
<td>Cooperative</td>
<td>Only ECA cooperates</td>
<td>Anger</td>
</tr>
<tr>
<td>Cooperative</td>
<td>Only human cooperates</td>
<td>Neutral</td>
</tr>
<tr>
<td>Competitive</td>
<td>Both defect</td>
<td>Neutral</td>
</tr>
<tr>
<td>Competitive</td>
<td>Both cooperate</td>
<td>Neutral</td>
</tr>
<tr>
<td>Competitive</td>
<td>Only ECA cooperates</td>
<td>Anger</td>
</tr>
<tr>
<td>Competitive</td>
<td>Only human cooperates</td>
<td>Joy</td>
</tr>
<tr>
<td>Competitive</td>
<td>Both defect</td>
<td>Neutral</td>
</tr>
</tbody>
</table>
While playing, participants will continuously encounter the ECA having a neutral emotion, it will only show Joy or Anger right after the outcome. Figure 2 shows the respective emotions at their highest intensity during the animation. To enforce participants to pay attention to the emotion there is a deactivation of the action buttons while the ECA is reacting to the outcome of the game.

3.3 Questionnaire
The post-game questionnaire consisted of a few questions, where participants were asked to rank a couple of attributes of the ECA on a scale of 1 to 5. These attributes are (Translated to English):

Fake – Lifelike
This property is to assess whether the participant felt like he was playing against a character that has some human qualities about it, or if the ECA was perceived as an image and nothing more.

Lying – Honest
This is to find out if people caught on to the fact that sometimes the ECA showed emotions that would indicate that it was cooperative, but it cooperated relatively little.

Competitive – Cooperative
This property was examined to see if participants notice the difference between ECAs that cooperate relatively often or not.

Emotionless – Full of emotion
This shows how much the participants registered the emotions of the ECA, or if they thought them to be of no value.

4. RESULTS
4.1 Participants
The survey was taken by 42 participants, of which 33 were male and the other nine female. The average age was 25.02 years, reflecting the fact that almost all participants were students at the University of Twente.

Initially, participants were split up into six groups dependent on the strategy and emotion type they had played against. However, since the neutral strategy type behaved exactly the same whether it displayed conflicting or consistent emotions, the respective two groups were merged. This results in a neutral group that is larger than the other groups. It also skews already skewed balance of the groups quite a bit.

4.2 Cooperation Rates
The design resulted in five groups of participants, the names of which are derived from the strategy and the emotion type of the computer opponent.

The groups were:
- Competitive-Consistent
- Competitive-Conflicting
- Neutral
- Cooperative-Consistent
- Cooperative-Conflicting

Following de Melo et. al. [5], cooperation rates were used as a measure to see how the participant responded to the ECA. An overview of cooperation rates for groups and combined groups is shown in Table 5.

Because none of the groups where distributed normally, the significant difference in the last column is calculated as a difference from the neutral control group using the non-parametric Mann-Whitney U Test, it is presented in the form: (U / 2-tailed significance).

Table 5. Cooperation Rates results for All Groups

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>Mean</th>
<th>SD</th>
<th>Sig</th>
</tr>
</thead>
<tbody>
<tr>
<td>All</td>
<td>42</td>
<td>0.2410</td>
<td>0.21340</td>
<td>-</td>
</tr>
<tr>
<td>Competitive-Consistent</td>
<td>6</td>
<td>0.16</td>
<td>0.15179</td>
<td>(32 / 0.235)</td>
</tr>
<tr>
<td>Competitive-Conflicting</td>
<td>4</td>
<td>0.16</td>
<td>0.16330</td>
<td>(21 / 0.296)</td>
</tr>
<tr>
<td>Neutral</td>
<td>16</td>
<td>0.25</td>
<td>0.16232</td>
<td>-</td>
</tr>
<tr>
<td>Cooperative-Consistent</td>
<td>7</td>
<td>0.3143</td>
<td>0.37053</td>
<td>(54 / 0.893)</td>
</tr>
<tr>
<td>Cooperative-Conflicting</td>
<td>9</td>
<td>0.2578</td>
<td>0.20795</td>
<td>(72 / 1.000)</td>
</tr>
</tbody>
</table>

The effect of the two independent variables (strategy and emotion type) was also calculated. Table 6 shows an overview similar to Table 5, describing the effect the different strategies had compared to the neutral group.

Table 6. Cooperation Rates results for Strategies

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>Mean</th>
<th>SD</th>
<th>Sig</th>
</tr>
</thead>
<tbody>
<tr>
<td>Competitive</td>
<td>10</td>
<td>0.16</td>
<td>0.14727</td>
<td>(53 / 0.151)</td>
</tr>
<tr>
<td>Neutral</td>
<td>16</td>
<td>0.25</td>
<td>0.16232</td>
<td>-</td>
</tr>
<tr>
<td>Cooperative</td>
<td>16</td>
<td>0.2825</td>
<td>0.28075</td>
<td>(126 / 0.940)</td>
</tr>
</tbody>
</table>

In the same way, we can break down the results for the conflicting and consistent emotions with respect to the neutral group, the same test as with the previous results was done. Table 7 shows an overview of these tests.

Table 7. Cooperation Rates results for Emotion Types

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>Mean</th>
<th>SD</th>
<th>Sig</th>
</tr>
</thead>
<tbody>
<tr>
<td>Consistent</td>
<td>13</td>
<td>0.2431</td>
<td>0.29096</td>
<td>(90 / 0.537)</td>
</tr>
<tr>
<td>Neutral</td>
<td>16</td>
<td>0.25</td>
<td>0.16232</td>
<td>-</td>
</tr>
<tr>
<td>Conflicting</td>
<td>13</td>
<td>0.2277</td>
<td>0.19417</td>
<td>(93 / 0.627)</td>
</tr>
</tbody>
</table>

4.3 Questionnaire
Afterwards participants were asked to fill in a questionnaire about the game they had just played. The questions concerned the characteristics of the ECA they had as their opponent.

Table 8. Answers to Questionnaire

<table>
<thead>
<tr>
<th>Question</th>
<th>N</th>
<th>Mean</th>
<th>SD</th>
<th>Median</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fake/Lifelike</td>
<td>39</td>
<td>2.10</td>
<td>0.995</td>
<td>2</td>
</tr>
<tr>
<td>Lying/Honest</td>
<td>39</td>
<td>2.85</td>
<td>0.904</td>
<td>3</td>
</tr>
<tr>
<td>Competitive/Cooperative</td>
<td>39</td>
<td>2.69</td>
<td>1.080</td>
<td>3</td>
</tr>
<tr>
<td>Emotionless/Full of</td>
<td>39</td>
<td>2.54</td>
<td>1.114</td>
<td>3</td>
</tr>
</tbody>
</table>

There were four questions, each about ranking the ECA on a scale from one to five for a specific characteristic. Averages and medians for each question are displayed in Table 8.
A more specific overview can be created as well. As the emotions that were actually shown by the ECA were recorded, we can look at the answers to the questions, and how they correlate to the emotions shown by the ECA. Table 9 shows how the number of shown emotions of a specific type correlates with the question answers.

Table 9. Correlation between number of Shown Emotions and Question Answers

<table>
<thead>
<tr>
<th>Shown Emotion Question</th>
<th>Joy</th>
<th>Neutral</th>
<th>Anger</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fake/Lifelike</td>
<td>0.123 (0.457)</td>
<td>-0.027 (0.868)</td>
<td>-0.025 (0.881)</td>
</tr>
<tr>
<td>Lying/Honest</td>
<td>-0.118 (0.467)</td>
<td>0.170 (0.301)</td>
<td>-0.144 (0.380)</td>
</tr>
<tr>
<td>Competitive/Cooperative</td>
<td>0.088 (0.596)</td>
<td>-0.083 (0.614)</td>
<td>0.057 (0.730)</td>
</tr>
<tr>
<td>Emotionless/Full of emotion</td>
<td>0.464 (0.003)</td>
<td>-0.123 (0.457)</td>
<td>-0.071 (0.666)</td>
</tr>
</tbody>
</table>

Table 10 shows how the number of cooperations by the ECA correlates to the answers of the questionnaire. The correlation coefficients in Tables 9 and 10 were calculated with the Pearson Correlation Coefficient method, the correlations are displayed in the form: correlation (significance 2-tailed).

Table 10. Correlation between Number of Cooperations and Question Answers

<table>
<thead>
<tr>
<th>Question</th>
<th>Correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fake/Lifelike</td>
<td>0.051 (0.822)</td>
</tr>
<tr>
<td>Lying/Honest</td>
<td>0.340 (0.121)</td>
</tr>
<tr>
<td>Competitive/Cooperative</td>
<td>0.376 (0.085)</td>
</tr>
<tr>
<td>Emotionless/Full of emotion</td>
<td>-0.005 (0.981)</td>
</tr>
</tbody>
</table>

4.4 Data filtering

One thing that stands out is the fact that quite a bit of the participants cooperated or defected without any change in behavior. These unconditional defectors and cooperators are of course valid strategies in real life, but it might be evident for any human being that an unconditional strategy that doesn’t take into account the fact that the opponent has a strategy of its own is far from productive in reaching the goal of reaching the highest combined score, as one always is vulnerable to exploitation when using unconditional cooperation, and shuts out opportunities for cooperation otherwise.

In addition, some of the feedback that was given by the participants in general conversation indicated that they might not have understood the goals of the game, as they laughingly said that beating the opponent was easy because they only had to choose Project Blue each time. This suggests they missed the point about cooperating with the ECA to obtain the mutual highest score.

A filter is therefore proposed, that removes any subjects that are so called ‘unconditionals’. They are defined as players that chose the same action in 23 or more of the 25 games. Figure 3 shows the histogram of the cooperation rates for the entire subject group. Figure 4 shows the cooperation rates for the filtered group, which is without the ‘unconditionals’. Both figures have their closest fitting normal distribution overlaid.

The proposed filter removes 18 of the 42 participants from the dataset: 8 participants that chose the same alternative all 25 times, 4 that deviated once and 5 that chose the other option two times. All other subjects were between a 5-20 split and a 12-13 split.

Table 11. Comparison of filtered datasets

<table>
<thead>
<tr>
<th>Dataset</th>
<th>N</th>
<th>Mean</th>
<th>SD</th>
<th>Skewn.</th>
<th>Kurtosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>No Filter</td>
<td>42</td>
<td>0.2410</td>
<td>0.21340</td>
<td>1.020</td>
<td>2.151</td>
</tr>
<tr>
<td>Filter</td>
<td>24</td>
<td>0.3567</td>
<td>0.09649</td>
<td>0.187</td>
<td>-1.149</td>
</tr>
</tbody>
</table>

Furthermore, Figure 5 shows a boxplot of the cooperation rates of the filtered dataset split over the Strategies of the ECA. The same analysis on the dataset split over the emotion
type of the ECA was done, but it showed the same results as the full dataset.

![Boxplot of filtered Cooperation Rates](image)

**Figure 5. Boxplot of filtered Cooperation Rates**
The discussion section will further elaborate on what this selection of data might mean for the results, and will offer inspirations for future work.

5. DISCUSSION

5.1 Experiment
This experiment was part of a larger group of experiments that were part of the Human Media Interaction group at the University of Twente. This had its advantages as well as its drawbacks.

Running the experiment (and initial research) in parallel with other students greatly improved upon the research as tips and tricks were exchanged. Sadly, the fact that it was a joint experiment also made it susceptible to problems over which there was little or no control.

In part this meant that the experiment was active late, which in its turn contributed to the low number of participants. Furthermore, because of the short time period, little to no testing could be done with the experiment, and it thus went online having a few problems. These were of course fixed, but it still meant that some subjects had already participated, which in turn meant that their data was not useful as the experiment had changed after it was collected.

Ideally the experiment would have used Elckerlyc [7], a tool that would have transformed neat Behavior Markup Language into movies of an ECA. Partly due to the complexity of this tool, but mostly due to the time constraints, this idea was abandoned in favor of the New York Universities “Face Demo” [9], which does not use exact markup, but it does allow for rapid prototyping and building of animations.

This means that there is no exact definition of the timing of the animation (in the form of a BML file), which hinders reproducibility. Another point is the fact that Elckerlyc produces much more lifelike animations, the Online Face Demo offers very recognizable emotions, but they can hardly be called realistic.

5.2 Results
Regrettably, the results show no significant differences between the researched groups. This means that this research cannot provide any proof of participants being affected by emotions of an ECA in the IPD.

Means for the cooperative group are higher than for the neutral group, which in its turn are higher than the competitive group (Table 6). Note that this ordering of groups agrees with the research done by de Melo et. al. [5] which tells us that persons are more likely to cooperate with an agent that is cooperative itself.

Although Table 7 shows that on itself the emotion type of an ECA seems to have no overall effect on the cooperation rate, Table 5 suggests that there is a difference when it is paired with a Competitive strategy, a difference that does not occur with a Cooperative strategy. This would be consistent with the theory of evolution, as detection of a lying cooperator has no real benefits, but failing to detect a lying competitive opponent could have severe consequences.

So even though no statistical significance can be obtained to prove these statements, they nonetheless are consistent with expected behavior.

Some interesting data comes from the questionnaire. Participants tend to rate the ECA as somewhat more Fake than Lifelike. This is understandable as the ECA is built using a simple polygon mesh. Participants also have no strong favor for the ECA Lying or it being Truthful, the same is true for the distinction between Cooperative and Competitive. Both are to be expected as they both occurred in equal amounts in the experiment.

There is no difference between two matching subgroups which have matching strategy and different emotion types, this is supported by the fact that participants rate the ECA as more Emotionless than Full of Emotion.

Another interesting point of data is seen in Table 9. The number of times a participant has been shown the Joy emotion correlates significantly with that person finding the ECA to be full of emotion. This correlation is absent with the Anger emotion. An underlying cause of the results could be the fact that the Joy emotion seems to be noticed much more than the Anger emotion.

Table 10 also shows some near-significant results about the subject’s perception of the ECA. Specifically, subjects do see the difference between a Competitive and Cooperative ECA. Furthermore, the Cooperative ECA is seen as more honest than is Competitive counterpart. This might be because participants can have feelings of spite against an uncooperative ECA.

5.3 Data filtering
Based on comments made by participants it was also interesting to look at the data with a filter applied to it. Although this data has less statistical meaning than the full dataset, it might help with understanding the rest of the data better.

Looking at individual groups has relatively little meaning with the filtered data, as they have become too small to make any reliable claim about them.

Figure 5 does show something that would be in line with previous research: people tend to cooperate more often with a cooperative opponent, and less with a competitive opponent.

This filtered data also has a more consistent view overall with prior comparable research by de Melo et. al. [5], in terms of average cooperation rates and standard deviations.
6. CONCLUSION
Now that we have discussed the results, a conclusion can be formed and we can try to answer the research questions. First, we will answer the sub questions, and from that we will come to an answer to the main research question and a conclusion on the entire research. Finally we will suggest possibilities for further research.

6.1 Research Questions
In section two we posed the following research questions. We can now answer them.

1.1. Do cooperation rates of humans in the IPD differ between ECAs with emotions consistent with their strategy and ECAs with conflicting emotions?
The research shows no evidence that cooperation rates differ between subjects playing against an ECA with consistent emotions and one with conflicting emotions. However, it does hint towards this being true, specifically for cooperative ECAs.

1.2. Do humans actively recognize ECAs with conflicting emotions?
Although it appears that humans do recognize emotions, there seemed to be no significant evidence that humans actively recognize conflicting emotions on an ECA.

1.3. Are the answers to questions 1.1 and 1.2 true for both cooperative and competitive ECAs?
There was a distinct difference in cooperation rates between both types of Strategy and with the competitive ECA there also seemed to be a slight difference between the emotion types. No real difference was found between competitive an cooperative ECAs when looking at the questionnaire answers.
With the answers to these three sub questions we can now answer the main research question.

1. Can humans correctly infer an agent’s strategy when his actions and emotions are conflicting?
The research shows no significant evidence that would support the claim that human subjects can or will make such a distinction, or even report about seeing a difference afterwards. Although some data does suggest humans doing so, specifically with cooperative ECAs.

It seems clear that neither one of the aspects is solely used in inferring an agent’s strategy, and that it is always a combination of the both.

6.2 Future Research
Even though the results of this research show no significant evidence that humans react to the ECA in the predicted way, there are some clues that would warrant further research.

Previous research [6] on the subject of humans interacting with ECAs, show humans responding and behaving differently to ECAs showing certain emotions. Although current research was not able to reproduce these findings with any significance, the results do tend to lean towards agreeing with that research.

One of the expected reasons for the lack of results is that the research was done in a relatively short timeframe, and as a consequence had a small set of participants. Furthermore, through the course of the research, it became somewhat apparent that the goals of the research were not made clear enough for some participants to fully understand them.

Another problem was the fact that the subject groups were not balanced enough, this in combination with the small size of the group of participants allowed for groups too small to effectively result in significant data.

The results also showed that the emotional response of the ECAs were not obviously recognized as such, with the exception of joy, better software for creating ECA movies can be used for this, for example Elckerlyc.

The data filtration showed some hints that the main research question might, in the end, be answered affirmatively, the questionnaire also hinted in the same direction.

It is therefore recommended that any future research should address these points. Although it seems that this research did not find any relation between the researched groups, a further, more carefully executed research, with more resources in terms of time to develop a more accurate application, and time and money to find a larger and more diverse subject group, could show that the questions posed in this paper can be answered with confirmation.

Research about ECAs and emotions, and interaction between humans and ECA, is something that will continue to develop into the future. As said in the introduction, lying is one of the most complex parts of the human communication, and research into the complexities of lying therefore deserves the attention of researchers.

7. REFERENCES
[7] Elckerlyc - A BML Realizer for continuous, multimodal interaction with a Virtual Human


[9] New York University Media Research Lab,
