Improving Humanness of Virtual Agents and Users’ Cooperation Through Emotions

Moojan Ghafurian, Neil Budnarain, and Jesse Hoey

Abstract—In this article, we analyze the performance of an agent developed according to a well-accepted appraisal theory of human emotion with respect to how it modulates play in the context of a social dilemma. We ask if the agent will be capable of generating interactions that are considered to be more human-like than machine-like. We conducted an experiment with 117 participants and show how participants rated our agent on dimensions of human-uniqueness (separating humans from animals) and human-nature (separating humans from machines). We show that our appraisal theoretic agent is perceived to be more human-like than the baseline models, by significantly improving both human-nature and human-uniqueness aspects of the intelligent agent. We also show that perception of humanness positively affects enjoyment and cooperation in the social dilemma, and discuss consequences for the task duration recall.

Index Terms—Affective computing, human likeness, OCC, prisoner’s dilemma, emotional intelligence, time perception, enjoyment

1 INTRODUCTION

In this paper, we investigate using emotions for generating artificially intelligent agents that seem more human-like. We argue that Artificial Intelligence (AI) research has mostly focused on building intelligence based on individual attributes that non-human animals do not possess, but that machines inherently do possess.

However, human intelligence also requires emotional attributes and social support, attributes that machines do not possess, whereas non-human animals do. The concept of “human-ness” has seen much debate in social psychology, particularly in relation to work on stereotypes and de-humanization or infra-humanisation [1]. In a comprehensive series of experiments, Haslam et al. [2] examined how people judge others as human or non-human (dehumanized). In their model, humanness is broken down into two factors. First, Human Uniqueness (HU) distinguishes humans from animals (but not necessarily from machines). Second, Human Nature (HN) distinguishes humans from machines (but not necessarily from animals). Human uniqueness traits are civility, refinement, moral sensibility, rationality, and maturity, as opposed to lack of culture, coarseness, amorality, irrationality, and childlikeness. Human nature traits are emotionality, warmth, openness, agency (individuality), and depth, as opposed to inertness, coldness, rigidity, passivity, and superficiality. Thus, while one can imagine both humans and animals having human uniqueness traits, animals would not tend to have these (they are considered to be coarse, amoral, etc). Similarly, while one can imagine both humans and animals having human nature traits, machines would not tend to have these (they are inert, cold, rigid, etc). While much research in AI is trying to build machines with HU traits (thus separating AI from animals), there is much less work on trying to build machines with HN traits (thus separating AI from machines). While both problems present challenges, the former problem is already “solved” to a certain degree by simply having a machine in the first place, as humans and machines are on the same side of the human uniqueness divide anyway. The latter problem is more challenging as the human nature dimension is exactly the dimension on which machines differ most from humans.

Several studies have verified that humanness and emotions of virtual agents can affect people’s behaviour and strategies [3], [4]. For example, Chowanda et al. [3] captured players’ emotions through their facial expressions and showed that non-player characters that have personalities and are capable of perceiving emotions can enhance players’ experience in the game. Further, Nonverbal behavior such as body gesture and gaze direction affects perception of cooperativeness of an agent [5].

In this paper, we use an appraisal-based emotional model in the same spirit as EMA [6], where emotional displays are made using the Ortony, Clore and Collins (OCC) model [7], and a set of coping rules are implemented to map the game history augmented with emotional appraisals to actions for the virtual agent. We refer to the agent based on this model as the OCC agent.

The OCC agent uses emotions to generate expectations about future actions [8], [9]. That is, it sees emotions as being related to an agent’s assessments of what is going to happen next, both within and without the agent’s control. The OCC agent computes expectations with respect to the
denotative meaning (or causal interpretation [6]) of the situation, and these expectations are mapped to emotion labels. The generated emotions are then used with a set of coping rules to adjust future actions.

We present results from a study involving $N = 117$ participants who played a simple social dilemma game with a virtual agent named “Aria”, whose emotional displays varied depending on the experimental condition. The game was a variation of Prisoner’s Dilemma (PD), in which each player could either give two coins or take one coin from a common pile. Players could maximize their returns by defecting while their partner cooperated, and although the Nash equilibrium is mutual defection, the players can jointly maximize their scores through mutual cooperation. The participants were awarded a bonus according only to their total score in the game, and so had incentive to cooperate as much as possible. Participants played a series of 25 games in a row, and then answered a questionnaire on how they felt about Aria on dimensions of human-uniqueness and human-nature taken from [2].

The participants were evenly split into three conditions that differed only in the emotional displays. The virtual human, Aria, displayed facial expressions and uttered canned sentences that were consistent with the game context and the emotional state given the condition. One condition was based on the OCC model, while two were baselines, one with randomly selected emotions, and one with no emotions. Our hypothesis was that the OCC agent would show more human-nature traits than the baselines, which can affect participants’ behaviour in the game and their enjoyment.

The primary contribution of this paper is to evaluate how appraised emotions relate to two important dimensions of humanness, and to investigate the impact of appropriate emotion modeling on perceived humanness of a virtual agent and users’ cooperation. Secondary contributions are a complete description of the prisoner’s dilemma in terms of OCC emotions and a demonstration in a simple environment.

2 RELATED WORK

Affective computing (AC) has formed as a sub-discipline of artificial intelligence seeking to understand how human emotions can be computationally modeled and implemented in virtual agents. While much current work is focused on social signal processing, the emphasis is on the detection, modelling, and generation of signals relating to social interactions without considering the control mechanisms underlying the function of emotion [10]. In a broad survey [11], Reisenzein et al. define emotional functions as being informational, attentional, and motivational, but point to a lack of explicit mechanisms for computational implementation.

Much of the work in AC on the function of emotions has focused on appraisal theories of emotion, as these give clear rules mapping denotative states to emotions and show a clear path for implementation, e.g., Ortony, Clore and Collins (OCC) [7] and Scherer’s component process model [12]. Efforts at integrating appraisal models in artificial agents started with Elliott’s use of an OCC model augmented with “Love”, “Hate”, and “Jealousy” to make predictions about humans’ emotional ratings of semantically ambiguous storylines and to drive a virtual character [13]. This was followed by the work of Gratch [6], [14], and OCC models were integrated with probabilistic models for intelligent tutoring applications in [15], [16]. A general-purpose game engine for adding emotions is described in [17].

Emotions have also seen a lot of attention for developing social robots and socially intelligent agents [18], [19], [20], [21], [22], [23], [24], as affective expressions are considered to be essential for developing emotionally/socially intelligent agents, along with other factors such as responsiveness to social clues [25]. For example, walker et al. (1994) [26] used a talking face on a computer screen and asked participants to to respond to a questionnaire. They showed that in the presence of a talking face, participants spent more time on the questionnaire (writing more comments and making fewer mistakes) as compared to those who saw the questions displayed by text on the screen. Similarly, a stern face yielded better results as compared to a neutral face [26]. This study showed that facial expressions shown by a virtual face can increase engagement [26]. Further, in the context of a competitive game (i.e., a competitive cards game called Skip-Bo), it was suggested that empathic feedback of an opponent agent can improve acceptance of the agent, but negative agent emotions may be beneficial and otherwise it may be irritating for the participants and induce stress [18]. Also, Demeure, Niewiadomski, and Pelachaud (2011) have shown that emotions can increase the perceived believability of an agent (which was also correlated with the agent being warm and competent), but suggested that believability may be different from being human-like [18]. Affective expressions can also affect people’s decision-making and perception of agents [3], [4]. Further, emotions can be effective approaches for communication, as they are intuitive to people and do not require much mental workload [27].

Furthermore, Ma et al. (2020) reviewed the empathic dialogue systems, identifying emotion-awareness, personality-awareness, and knowledge accessibility as the key aspects of such systems, which could enhance perception of the system. While there were multiple studies exploring each of these factors, many challenges were discussed that showed that we are still at very early stages of creating dialogue systems with real human-human interaction capabilities [28]. One of these challenges, which was proposed as an open research problem, was the ability of the system to use emotions that also comply with a goal [28].

To enable virtual agents with the ability to show emotions, many methods have been proposed for creating facial expressions for virtual agents, etc. (e.g., [29], [30], [31]), and to understand users’ emotions, different methods for sentiment analysis have been proposed to detect emotions and sentiments from different sources, such as video, text, dialogue, etc. (e.g., [32], [33], [34]). Further, the effect of emotions on decision-making has been studied and different approaches have been proposed, many of which borrow from behavioural economics and cognitive science to characterize the consequences of emotion in decision making and integrate this knowledge as “coping rules” [6], “affect heuristics” [35], or short-circuit “impulsive behaviours” [36] to characterize or influence agents’ behaviours.

Investigations into the role of emotions in modifying behaviours in a PD have looked at how disappointment and anger can be used to promote forgiveness and retaliation,
In each round, the players' decisions are hidden from each other. In the game, Aria, the virtual opponent, is sitting in front of the participants. On the right, there is a large pile of coins that the players use. On the left there are two piles of coins, one pile representing the coins that Aria has received so far and the other showing human player's coins. Upon selecting both action and emotion, the results will be revealed (coins will appear on the table and will move to the players’ piles) and both players will see the other player’s emotion. Aria’s emotions are shown in her facial expressions and utterances, and players’ emotions are shown via the selected emoji.

3.1 Emotion Dictionary

We used a set of 20 emotions compiled from the OCC model (see Section 4) and these were mapped to a three dimensional emotion space with dimensions of Evaluation (E), Potency (P), and Activity (A), which is sometimes known as the VAD model where the terms are Valence (E), Activity (A), and Dominance (P) [46]. We used an emotion dictionary consisting of a set of 300 emotion words rated on E, P, and A by 1027 undergraduates at the University of Indiana in 2002-2004 [47]. We manually found the same or a synonymous word in the dictionary for each of the 20 OCC emotion words. We also selected a set of 20 emojis, one per OCC emotion, to be used in the game play as described above.

An important emotion in PD is regret. If a player defects, but regrets, it is quite different than when a player defects but shows no regret. In the following, we define regret as any of the four OCC emotions “remorse”, “distress”, “shame”, or “fears-confirmed”. Only this last term does not have a direct equivalent in the Indiana dictionary, and for it we found the term “heavy hearted” (E: -1.03; P: -0.55; A: -1.15).

3.2 Facial Expressions

While there are many existing methods to generate facial expressions (e.g., [29], [30], [31]), we did not have the flexibility to change Aria’s action units and changing Aria’s expressions was limited to using a combination of the “HSF” space: (1) Happy/Sad, (2) Surprise/Anger, and (3) Fear/Disgust. Therefore, Aria’s facial expressions were generated with three controls that map to specific sets of facial muscles. A setting of these three controls yields a specific facial expression by virtually moving the action units in the face corresponding to that emotion by an amount proportional to the control. For example, “happy” is expressed with AU6 + AU12: cheek raiser and lip corner puller. In general, although the virtual humans’ face can be controlled by moving individual muscles, like the inner eyebrow raise, groups of these are highly correlated and move in recognizable patterns. Therefore, these three dimensions of musculature movement could be sufficient.

To map an emotion label from our set of 20 to a facial expression, first the emotion ratings in the EPA space were found using the Indiana dictionary. Afterwards, the distance to each end-point of the HSF space was computed using the EPA ratings shown below, also from the Indiana dictionary, and these distances were used to set the HSF controls directly. Fig. 2 shows a few examples of the expressions.
Aria had a normal “quiescent” state in which she blinked and slightly moved her head from side to side in a somewhat random way. The emotions were applied for 10.5 seconds and Aria’s face smoothly transitioned to a weaker representation of the same emotion and the quiescent state.

### 3.3 Speech

Depending on the emotion being shown, Aria also provided a verbal comment about the feeling. This was added to ensure that the participants will understand the emotions even if they could not identify the facial expressions. For each emotion, the sentences were selected from a predefined set of four, one for each combination of agent action, human action, and binary indicator of valence (E). Speaking and facial expressions were possible at the same time. Lip movement during speech was based on a proprietary algorithm. To define the sentences, an embedding (vector) for each emotion label was computed using the pre-trained Word2Vec model, which was trained on part of the Google News dataset where the model contains 3 million words and 300 dimensional vectors [48]. Phrases were then embedded as follows: (a) stop words were removed using the stop word list provided by the NLTK library for the English language, (b) the embedding of a phrase was then computed by taking the mean of the Google embeddings of all the words in a phrase; and (c) given an emotion label, the closest phrase was queried by computing the cosine similarity (dot product) of the vector representing the emotion label with all of the phrase vectors.

### 4 OCC Model of Emotion

According to the OCC model [49], emotions arise as a valenced reaction to the consequences of events, to the actions of agents, and to the aspects of objects. In our game situation, the aspects of objects (leading to the emotions of love and hate) do not change and so may influence overall mood but will not change substantially over the course of the interaction. We therefore focus on actions and events only. Emotions in these categories are caused by the immediate payoffs, or by payoffs looking into the future and past. Within each category, there are a number of further distinctions, such as whether focus is on the self or the other, and whether the event is positive or negative. There are 20 emotions in the model after removing “love” and “hate”.

### 4.1 OCC Appraisals in the Prisoner’s Dilemma

Table 1 shows the OCC interpretation of emotions in our Prisoner’s Dilemma game. Each row shows a situation that may happen in the game. The “Game Play” column gives the most recent move of both agent (Aria) and player (human), and the player’s emotion in the previous round. The columns on the right (under “Appraised Emotions”) show the momentary emotions appraised on the consequences of events and on the actions of agents, both of which are appraised for both self and other, as well as the prospect-based consequences of events, which are evaluated on each subsequent turn. Emotional intensity is not modeled, but could be added to increase realism. The middle column (“Valenced Appraisals”) shows the process of generating the suggested emotions through the OCC model.

For example, in the third row in Table 1, we see a situation where the player has cooperated (given 2) in the previous round, the player cooperates again and shows a positive emotion, but Aria defects by taking one. In this situation, Aria will be happy as a result of the outcome. If Aria focuses on the momentary appraised emotion, based on these actions and the OCC interpretation of emotions, her emotions will be different depending on whether she thinks about the consequence of her actions for self, or for the other (the player). If Aria thinks of the consequence of the actions for self, she will feel joy, admiration, and gratitude. However, thinking about the consequence for the other, Aria will feel shame (disapproving her own action as she has done something wrong, since the other person has cooperated and shown a positive emotion).

Prospect-based emotions arise because Aria looks into the future and predicts how things will evolve. Focusing on the future, Aria will feel hope and satisfaction in this situation (because players’ action and emotion are promising and she assumes that things will go well).

### 4.2 Coping

Once an emotion is appraised, coping is used to figure out what action to take. Five coping strategies are taken from [6]: acceptance, seeking support, restraint, growth, and denial, and these are applied as shown in Table 2. At the game’s start, Aria’s hope leads to the support seeking coping mechanism, and thus to an initial cooperative action.

### 5 Experiment 1

To assess how humanness of the OCC agent is perceived, we ran an experiment on Mechanical Turk, where the participants played our Prisoner’s Dilemma game against different agents with different strategies and emotional
displays. We then asked participants to evaluate each agent on Human Nature and Human Uniqueness traits. In the following, we will present our methodology and results.

### 5.1 Method

The experiment consisted of two parts: the Prisoner’s Dilemma game and a questionnaire. Participants played 25 rounds of the game against an agent, which was randomly assigned based on the experimental condition. Afterwards, they filled out a questionnaire assessing how they perceived different aspects of humanness of the agent that they were assigned to. We ensured that the participants would pay attention and try to maximize their outcome by providing a bonus according to the points they earned in the game. The amount of the bonus was significantly larger than the initial payment. Further, participants were told that they will play up to 30 rounds because knowing the number of rounds can affect people’s strategy in the final rounds.

#### 5.1.1 Hypotheses

The hypotheses of this experiment were as follows.

---

### TABLE 1

| Previous | Most Recent | Emotion | Consequences | Action of agents | APPRAISED EMOTIONS |
|----------|-------------|---------|--------------|-----------------|-------------------|
| Player   | Aria        | Player  |              |                 | Momentary          | Prospect-Based    |
| give 2   | give 2      | give 2  | any          | yes             | happy-for         | hope satisfaction |
| take 1   | give 2      | give 2  | any          | yes             | happy-for         | hope relief       |
| give 2   | take 1      | give 2  | positive     | no              | pity              | hope satisfaction |
| take 1   | take 1      | give 2  | positive     | no              | pity              | hope relief       |
| give 2   | take 1      | give 2  | negative     | no              | gloat            | fear satisfaction |
| take 1   | take 1      | give 2  | negative     | no              | gloat            | fear relief       |
| give 2   | give 2      | take 1  | no regret    | no              | resentment        | fear disappointment|
| take 1   | give 2      | take 1  | no regret    | no              | resentment        | fear fears-confirmed|
| give 2   | give 2      | take 1  | regret       | no              | resentment        | hope disappointment|
| take 1   | give 2      | take 1  | regret       | no              | resentment        | hope fears-confirmed|
| take 1   | take 1      | take 1  | any          | no              | pity              | fear fears-confirmed|
| give 2   | take 1      | take 1  | any          | no              | pity              | fear disappointment|

The “consequences” and “actions of agents” correspond to the OCC decision tree. ▲ means “pleased”, ▼ means approving, ◇ means desirable, and □ means confirmed. Aria is ambivalent for all lines not shown. For example, in the case where Aria gives while the player takes but shows regret, Aria does not disapprove of the player’s action anymore (because he is showing regret), but does not actually approve of it either, so sits on the fence and does not feel admiration or reproach.
**TABLE 2**

Coping Strategies for the PD Bot, Including Last Player Emotion, and the Last Two Player Moves

| Player moves | Player strategy coping | example | next move |
|--------------|------------------------|---------|----------|
| t-2 | t-1 | emotion (t-1) | acceptance: live with bad outcome | “oh well, we’re doomed” | take 1 |
| take 1 | take 1 | - | acceptance: live with bad outcome | “oh well, we’re doomed” | take 1 |
| take 1 | give 2 | positive | growth: positive reinterpretation | “this might be turning around” | give 2 |
| take 1 | give 2 | negative | growth+denial: positive reinterpretation | “maybe he didn’t mean that emotion” | give 2 |
| give 2 | take 1 | regret | restraint: hold back negative, keep trying | “he’s a good person really” | give 2 |
| give 2 | take 1 | not regret | denial: deny reality, continue to believe | “maybe its not so bad” | give 2 |
| give 2 | give 2 | - | seek support: understanding and sympathy | “Let’s cooperate together on this” | give 2 |

- **H1**: Showing proper emotions, as in OCC model, could improve perception of human-likeness of intelligent agents.
- **H2**: Showing proper emotions, as in OCC model, could improve users’ cooperation with the agent.
- **H3**: Showing proper emotions will increase users’ enjoyment.
- **H4**: Interactions with the OCC and Random agents would be recalled longer, as they talked and showed emotional expressions during the game, and therefore presented participants with more information.

As the emotions are important elements of humans’ communication, we believe that making virtual agents emotionally intelligent will positively increase perception of human-likeness (H1), especially on the human nature aspects that were discussed before. Further, as emotions can increase perception of social presence [50], and social presence is shown to positively affect trust [51], we hypothesized that showing proper emotion can increase participants’ cooperation with the agent (H2). Also, the higher perception of social presence can increase users’ enjoyment (H3).

In a game with multiple rounds, Ghafurian, Reitter, and Ritter [52] showed that the participants who see faster changing/more information during delays between the game rounds (varied by changing the speed of a countdown) would recall the duration of the game longer. However, they showed that more information would reduce negative effects of impatience during the game (i.e., time would be felt to pass faster), leading to a better overall experience. This was explained by the model of memory markers [53], which proposes that the durations with more memory markers (of any visual or cognitive nature) would pass faster, whereas memory markers would be recalled longer, as people will refer to those memory markers when estimating the durations of past events. We believe that showing emotions can act in a similar manner, therefore, while the participants might feel passing of time as faster when interacting with agents that show emotions (also resulting in a higher enjoyment as hypothesized in H2), they will recall the duration of the game as longer afterwards (H4).

5.1.2 Experimental Conditions

The experiment had three between-participant conditions. In all conditions, participants played the same number of game rounds against an opponent (Aria) and answered the same survey. However, the behaviour and emotional displays of the agent changed depending on the experimental condition. The conditions of the experiment were as follows:

1) **OCC**: Agent acts according to Table 2, which is a tit-for-2-tats strategy. Emotional displays are from the set defined in Table 1. Among the set of appropriate emotions for each scenario, it will be randomly selected whether Aria should show an emotion that is prospect-based, or one that is momentary-based, and also, if momentary-based is selected, it will be randomly selected whether the consequence for self or for the other is taken into account (e.g., in the situation shown in the third row of Table 1, an emotion will be randomly selected from the set of [pity, join, admiration, gratitude, shame, hope, and satisfaction]). Facial expressions and sentences are applied as described above.

2) **Emotionless**: Agent plays tit-for-2-tats (cooperates immediately upon cooperation, but defects only after two defections), shows no emotional expressions in the face and says nothing. This agent still shows quiscent behaviours.

3) **Random**: Agent plays tit-for-2-tats. Emotional displays are randomly drawn from the set of 20 emotions, and facial expressions and sentences are selected on the basis of that.

The Emotionless agent is added to ensure that the participants are paying attention to the emotions when rating the humanness of the agents. The Random condition enables us to study whether the participants pay attention to the differences in the emotional displays, and the relationship between their actions/emotions and the agent’s emotions, when rating the humanness of the agents.

5.1.3 Questionnaire

We used four sets of questions before and after the game. These questions were as follows:

1) **Demographic Questionnaire**: Before the game, participants were asked to provide their demographic information (i.e., age and gender). We used this information to control for possible effects of gender and age on perception of the humanness of the agents. Participants could decide not to disclose this information.

2) **Humanness Questionnaire**: After the game, we used the humanness assessing questionnaire to assess participants’ perception of agents’ emotions and behaviours. The humanness questionnaire consisted of 18 questions. The first two questions asked participants to rate to what extent they thought that the agent behaved human-like/animal-like, and to what
extend they thought it behaved human-like/machine-like. The following 16 questions used the traits proposed by Haslam et al. [2] and assessed different Human Nature and Human Uniqueness traits in more details.

3) **Time Perception Question:** After the Humanness Questionnaire, participants were asked to estimate how long their interactions with the agent (i.e., the game) lasted through the verbal estimation method [54]. We used this question to study the effect of agent’s emotional displays on time perception to study H4.

4) **Enjoyment Question:** After the time perception question, participants were asked to rate how much they enjoyed playing the game. We used this question to study whether different emotional displays can affect participants’ satisfaction to study H3.

5) **IDAQ Questionnaire:** After answering all other questions, participants answered the IDAQ questionnaire proposed by Waytz et al. [55]. The results from this questionnaire was used to account for individual differences in the general tendency to anthropomorphize.

A continuous slider was used in all questions, except IDAQ, which uses an 11-scale: the standard scale for this questionnaire [55]. In addition to these questions, a total of six attention-check questions with clear answers (e.g., “How many ‘a’s are in the word “Aria”?”) were randomly placed among the questions in the Humanness and IDAQ questionnaires to ensure that the participants paid attention.

5.1.4 Procedure

Participants first signed the consent form and provided their demographic information. Then they played 25 rounds of the game against one of the agents, which was randomly assigned to them. After completing all 25 rounds of the game, participants answered the aforementioned set of questions: humanness, time perception, enjoyment, and IDAQ. Repeated participation was not allowed. We ensured that the participants saw and heard Aria, and used a browser that was compatible with our platform.

5.1.5 Participants

Participants were recruited on Amazon Mechanical Turk. 124 participants completed the game and the questionnaire (74 male, 48 female, 1 other, and 1 did not wish to share, age range: [21,74]). The data from 6 participants (4 male and 2 female) were removed as they failed to pass the attention checks. Data from 1 participant (male) was removed as he was not able to hear the agent properly. Participants received an initial payment of $0.7 and a bonus according to their performance in the game ($0.05 for each point they earned). Participation was limited to workers in US and Canada who had completed at least 50 HITs and had a prior MTurk approval rate of 96 percent. The experiment was approved by the University of Waterloo’s Research Ethics Board.

6 RESULTS - EXPERIMENT 1

In this section, we will first demonstrate how playing against different agents affected perception of humanness. We then show the effects on participants’ cooperation rates, enjoyment, and time perception.

6.1 H1 - Humanness

We assessed all the agents based on the rating of HN and HU traits. Fig. 3 shows the results. As hypothesized, the OCC model was perceived to be more human-like on both HN and HU aspects. We fit two linear mixed effect models predicting HN and HU ratings based on experimental condition. IDAQ, the general tendency to anthropomorphize, was controlled for. We also contolled for possible effects of age, gender, and bonus (as the final bonus may affect people’s perception of the agent). A random effect based on the day on which the experiment was run was fitted. The modelling results for HN and HU ratings are shown in Tables 3 and 4, respectively. The OCC agent’s HN traits were perceived to be significantly higher than the Emotionless agent, and its HU traits were perceived to be significantly higher than the Random agent. That is to say, overall, the OCC was perceived to be significantly more human-like as compared to the other two conditions. Also, these results suggested that while showing random emotions may help improve perception of HN traits, it can significantly and negatively reduce perception of HU traits, perhaps because
it would make the agent’s behaviour look inappropriate (e.g., by breaking social rules) or irrational. Further, the bonus has significantly affected perception of the HU traits, as these traits mostly describe perception of the agent’s actions (e.g., rationality, sensibility, etc.). However, we did not see any effect of bonus on perception of HN traits.

### 6.2 H2 - Cooperation

Next, we checked whether different emotional displays affected participants’ strategies. All agents played the same strategy (i.e., tit-for-two-tats); therefore, the difference in cooperation rates among conditions can reflect the effect of the different emotional displays on participants’ tendency to cooperate (in other words, trusting the agent). Fig. 4 shows the results. OCC has the highest cooperation rate and seems to encourage cooperation. This difference is significant between the OCC and Random agent (se = 1.851, t = -2.006, p < 0.05), however, we did not see a significant difference between the cooperation rates of the OCC and Emotionless agent.

### 6.3 H3 - Enjoyment

We know that perception of humanness of virtual agents can affect people’s enjoyment in games [3]. Here we ask what attributes of humans contribute to this effect. Therefore, we look into HN and HU traits independently, hypothesizing that HN traits are the key factors for enjoyment, as they distinguish humans from machines.

### 6.4 H4 - Time Perception

Fig. 6 shows the results for the game duration recall. Duration of the game was estimated to be significantly larger by those who played against OCC (se = 0.431, t = 2.635, p < .01) and random (se = 1.352, t = 3.100, p < .005) agents, as compared with the emotionless agent (significance was calculated through linear regressions).

While recalling a duration to be shorter might be an advantage in many situations, it could be beneficial in games and situations where the users are willing to spend more time with an agent.

### 7 Discussion

Affective experience increases engagement [56] of users, improves loyalty [57], and influences perception of humanness of the agents, which can affect people’s behaviour and enjoyment [3]. In this paper, we studied how emotions affect
perception of different dimensions of humanness of the agents and users’ trust, enjoyment, and time perception. We utilized Haslam et al.’s definition of humanness [2] to study how emotions affect perception of Human Nature (HN), distinguishing humans from machines, and Human Uniqueness (HU), distinguishing humans from animals. We asked how emotions affect peoples’ perception of HU and HN traits of an agent, and as a result, their opinion about, and behaviour towards the agents. We hypothesized that although there is an emphasis on improving the HU dimension of computers (as a result of making computers more refined, rational, and moral), improving the HN dimension has seen relatively limited attention. With emotionality being an aspect of HN, we hypothesize that agents capable of showing emotions will be perceived more human-like, especially on the HN traits.

We used a social dilemma, the prisoner’s dilemma, to test this hypothesis. In prisoner’s dilemma, the players cooperate if they trust the opponent, so this game enables us to study how emotions and perception of humanness of agents can affect trust. Although all the agents (opponents) played the same tit-for-two-tats strategy, the difference in their emotional displays significantly affected perception of the human-like traits. The OCC agent, capable of showing meaningful emotions, was perceived significantly more human-like on both HN and HU traits (H1 confirmed). This significantly improved participants’ cooperation rate and enjoyment (H2 and H3 confirmed). Any expression of emotion, even by the Random agent, improved perception of Human Nature traits (warmth, openness, emotionality, individuality, and depth). However, displaying random emotions negatively affected Human Uniqueness traits (civility, refinement, moral sensibility, rationality and maturity), as it can make the agent look irrational and immature. Therefore, while showing proper emotions enables computer agents with Human Nature traits and fills the gap between humans and machines, showing random emotions that are not necessarily meaningful is even worse than showing no emotions for the Human Uniqueness traits.

Furthermore, we showed that the presence of emotions can increase duration recall of the game. Thus, emotions can be beneficial in some tasks, where users’ recall of a longer interaction can positively affect their perception of helpfulness of the agent. However, for tasks that require a quick response from the agent, presenting emotions might not be as beneficial as it can reduce the perception of the efficiency of the agent.

The general anthropomorphism tendency (measured through the IDAQ questionnaire) significantly and negatively affected cooperation. This may suggest an uncanny valley effect: those who anthropomorphized more perceived Aria to be more similar to humans, which resulted in disliking Aria and not trusting her [58].

Finally, age significantly affected ratings of Human Nature traits. All the agents were perceived more human-like on the HN traits when age increased. This may be because the younger adults are more used to seeing avatars and characters in computer games, which look similar to humans, thus have a higher standard in mind regarding virtual agents. Another interesting observation was that the amount of bonus significantly affected perception of HU traits. Possibility because the participants believed that the agent (i.e., their opponent in the game) was in fact responsible for what they earned (the results), and associated a higher bonus to a better performance of the opponent (despite the fact that the strategies of all the agents were identical).

8 LIMITATIONS AND FUTURE WORK
Our work has a number of limitations. First, the OCC coping mechanism is theoretically difficult to justify and is usually specified in a rather ad-hoc manner [59]. In the simple game considered here, it provides a reasonable approximation and yields a strategy often used by humans (tit-for-two-tat). Similar coping strategies could be defined using an “intrinsic reward” generated by the appraisal variables. As reviewed by Broekens et al. [60], this intrinsic reward requires some weighting factor (e.g., φ in [60]) which is difficult to specify. In this simple game, we could, for example, consider that motivational relevance, which is inversely proportional to the distance from the goal, may be larger in cases where the agent predicts cooperation (e.g., the other player cooperates, or defects but shows regret), and smaller when the agent predicts defection (e.g., the other player defects and shows no regret, or defects multiple times). Motivational relevance would add intrinsic reward to the cooperation option, making it game theoretically optimal compared to defection. In the give2-take1 game, this would require adding a reward of 1 to the cooperation option when motivational relevance was high (when there was “hope”) and not doing so when motivational relevance was low (when there was “fear”).

A second limitation is that emotions are displayed in the face based only on a dimensional emotion model (EPA space), which neglects the semantic context. For example, while repellant and reverent have different meanings which should result in different facial expressions, their EPA ratings are almost identical. Therefore there are some instances where mapping from EPA → facial expressions does seem accurate, but emotion label → facial expression does not seem as accurate. Furthermore, the generated facial expressions were not tested and we cannot ensure that the participants understood all the facial expressions. However, the sentences accompanying these
facial expressions (e.g., “I am disappointed” for disappointed, “I am sorry” for ashamed, “thank you, I am delighted” for grateful, etc.) helped, to some extent, to ensure that the participants perceived a range of emotions.

A third limitation is the limited number of emojis used to allow human players to express emotions, and the interpretation given to those emojis by the participants. While the emotion word can be seen by hovering, a better method would involve facial expression recognition to extract emotional signals directly.

Lastly, we used a set of 20 emotions as defined in the OCC model, however, the same results might be obtained with a smaller set of emotions and using a different emotion model. Our emphasis in this study was on the use of meaningful emotions, and how this can affect perception of two dimensions of humanness. Future work would benefit from comparing different sets of emotions to study how perception of these two humanness dimensions are affected by the range of emotions that an agent shows. It can be also investigated if/how the results would be different if only prospect-based/momentary emotional displays were used.

9 Conclusion

This paper described our work towards understanding the effect of emotions on different dimensions of humanness of computer agents, as well as on users’ cooperation tendency, enjoyment, and time perception. We studied traits distinguishing humans from machines (Human Nature) and those distinguishing humans from animals (Human Uniqueness), and showed that proper expressions of emotions increases perception of human nature of agents. While researchers can successfully improve perception of Human Uniqueness traits by making agents smarter, emotions are critical for perception of Human Nature traits. This improvement also positively affected users’ cooperation with the agent and their enjoyment. Further, we showed that if emotions are not reflected properly (e.g., generated randomly), they can have negative effects on perception of humanness (HU traits) and can reduce the quality of social agents, even compared to when the agent does not reflect any emotions. Therefore, it is important to find models that accurately understand and express emotions, and utilize them properly in developing virtual agents, should those agents need to be perceived as more human-like.

Acknowledgments

This work was supported by Natural Sciences and Engineering Research Council of Canada (NSERC) and Social Sciences and Humanities Research Council of Canada (SSHRC) through the Trans-Atlantic Platform’s Digging into Data Program, and AGE-WELL, Inc., a Canadian Network of Centers of Excellence. The authors would like to thank Nattawut Ngampatipatpong and Sarel van Vuuren for help with the virtual human.

References

[1] S. Demoulin, J.-P. Leyens, M.-P. Paladino, R. Rodriguez-Torres, A. Rodriguez-Perez, and J. Dovidio, “Dimensions of “uniquely” and “non-uniquely” human emot.,” Cogn. Emot., vol. 18, no. 1, pp. 71–96, 2004.
[2] N. Haslam, S. Loughman, Y. Noorishad, and P. Bain, “Attributing and denying humanness to others,” Eur. Rev. Soc. Psychol., vol. 19, no. 1, pp. 55–85, 2008.
[3] A. Chowanda, M. Flintham, P. Blanchfield, and M. Valstar, “Playing with social and emotional game companions,” in Proc. Int. Conf. Hum. Comput. Interact., 2015, pp. 85–95.
[4] C. M. De Mol, P. Carnevale, and J. Gratch, “The influence of emotions in embodied agents on human decision-making,” in Proc. Int. Conf. Intell. Virt. Agents, 2010, pp. 357–370.
[5] C. Straßmann, A. R. von der Putten, R. Yaghoubzadeh, R. Kaminski, and N. Krämer, “The effect of an intelligent virtual agent’s nonverbal behavior with regard to dominance and cooperativeness,” in Proc. Int. Conf. Intell. Virt. Agents, 2016, pp. 15–28.
[6] J. Gratch and S. Marsella, “A domain-independent framework for modeling emotion,” Cogn. Syst. Res., vol. 5, no. 4, pp. 269–306, 2004. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1389041704000142
[7] A. Ortony, G. Clore, and A. Collins, The Cognitive Structure of Emotions. Cambridge, U.K.: Cambridge Univ. Press, 1988.
[8] N. H. Frijda, “The psychologists’ point of view,” in Handbook of Emotions, 3rd ed. New York, NY, USA: Guildford Press, 2010, pp. 68–77.
[9] R. Zajonc, “Feeling and thinking: Closing the debate over the independence of affect,” in Feeling and Thinking: The Role of Affect in Social Cognition. Studies in Emotion and Social Interaction, vol. 2, J. P. Forgas, Ed. New York, NY, USA: Cambridge Univ. Press, 2000, pp. 31–58.
[10] A. Vinciarelli et al., “Bridging the gap between social animal and unsocial machine: A survey of social signal processing,” IEEE Trans. Affective Comput., vol. 3, no. 1, pp. 69–87, First Quarter 2012.
[11] R. Reisenzein et al., “Computational modeling of emotion: Toward improving the inter-and intradisciplinary exchange,” IEEE Trans. Affective Comput., vol. 4, no. 3, pp. 246–266, Third Quarter 2013.
[12] K. R. Scherer, A. Schorr, and T. Johnstone, Appraisal Processes in Emotion. London, U.K.: Oxford Univ. Press, 2001.
[13] C. Elliott, “Hunting for the holy grail with “emotionally intelligent” virtual actors,” ACM SIGART Bull., vol. 1, no. 1, pp. 20–28, 1998–1999.
[14] J. Gratch, “Emilie: Marshelling passions in training and education,” in Proc. 4th Int. Conf. Autom. Agents, 2000, pp. 325–332.
[15] J. Sabourin, B. Mott, and J. C. Lester, “Modeling learner affect with theoretically grounded dynamic Bayesian networks,” in Proc. Affect. Comput. Intell. Interact., 2011, pp. 286–295.
[16] C. Conati and H. Maclaren, “Emotionally building and evaluating a probabilistic model of user affect,” User Model. User-Adapted Interact., vol. 19, pp. 267–303, 2009.
[17] A. Popescu, J. Broekens, and M. van Someren, “GAMYGDAL: An emotion engine for games,” IEEE Trans. Affective Comput., vol. 5, no. 1, pp. 32–44, First Quarter 2014.
[18] C. Becker, H. Prendinger, M. Ishizuka, and I. Wachsmuth, “Evaluating affective feedback of the 3D agent max in a competitive cards game,” in Proc. Int. Conf. Affect. Comput. Intell. Interact., 2005, pp. 466–475.
[19] I. Poggi, C. Pelachaud, F. de Rosis, V. Carofiglio, and B. De Carolis, “Greta. A believable embodied conversational agent,” in Multi-modal Intelligent Information Presentation. Berlin, Germany: Springer, 2005, pp. 3–25.
[20] J. D. Velásquez, “An emotion-based approach to robotics,” in Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst., 1999, pp. 225–240.
[21] J. Saldien, K. Goris, B. Vanderborght, J. Vanderfaeillie, and M. van Someren, “KASPAR–A minimally expressive humanoid robot for modeling emotion,” Cogn. Syst. Res., vol. 3, no. 1, pp. 69–87, First Quarter 2012.
[22] M. Ghafurian, G. Lakatos, Z. Tao, and K. Dautenhahn, “Design and evaluation of affective expressions of a zoomorphic robot,” in Proc. Int. Conf. Soc. Robot., 2020, pp. 1–12.
[23] K. Dautenhahn et al., “KASPAR—A minimally expressive humanoid robot for human–robot interaction research,” Appl. Bionics Biomechanics, vol. 6, no. 3/4, pp. 369–397, 2009.
[24] J. H. Walker, L. Sproull, and R. Subramani, “Using a human face in an interface,” in Proc. SIGCHI Conf. Hum. Factors Comput. Syst., 1994, pp. 85–91.
[25] B. Van Acker, D. Parmentier, P. Vierick, and J. Saldien, “Understanding mental workload: From a clarifying concept analysis toward an implementable framework,” Cogn. Technol. Work, vol. 20, pp. 351–363, 2018.
Moojan Ghafurian received the PhD degree in information sciences and technology from the Pennsylvania State University, State College, Pennsylvania, in 2017. She is currently a research assistant professor with the Department of Electrical and Computer Engineering, University of Waterloo. She was previously the inaugural Wes Graham postdoctoral fellow (2018-2020) and a research member at the David R. Cheriton School of Computer Science. Her interests and background are in human-computer interaction, human-robot interaction, affective computing, and social robotics. She is a member of the Social and Intelligent Robotics Lab (SIRRL) and Computational Health Informatics Lab (CHIL), University of Waterloo.

Neil Budnarin received the honours bachelor’s of computer science (BCS) and the master’s of mathematics degrees from the University of Waterloo, Waterloo, Canada. He is currently a research associate with the Computational Health Informatics Lab (CHIL). He has worked on the integration of an emotional virtual human into a handwashing system designed to assist with cognitive disabilities such as Alzheimer’s. His research interests include leveraging machine learning algorithms in dimensional models of emotions.

Jesse Hoey received the PhD degree in computer science from the University of British Columbia, Vancouver, Canada, in 2004. He is currently a professor with the David R. Cheriton School of Computer Science, University of Waterloo, where he leads the Computational Health Informatics Laboratory (CHIL). He is also an adjunct scientist with the Toronto Rehabilitation Institute and a faculty affiliate with the Vector Institute both in Toronto, Canada. He is an associate editor of the IEEE Transactions on Affective Computing and an area chair for the International Joint Conferences on Artificial Intelligence (IJCAI 2019–2021).

For more information on this or any other computing topic, please visit our Digital Library at www.computer.org/csdl.