Abstract—The behavior of virtual characters in computer games is usually determined solely by decision trees or finite state machines, which is detrimental to the characters’ believability. It has been argued that enhancing the virtual characters with emotions, personalities, and moods, may make their behavior more diverse and thus more believable. Most research in this direction is based on existing (socio-)psychological literature, but not tested in a suitable experimental setting where humans interact with the virtual characters.

In our research, we use a simplified version of the personality model of Ochs et al. [1], which we test in a game which has human participants interact with three agents with different personalities: an extraverted agent, a neurotic agent, and a neutral agent. The model only influences the agents’ emotions, which are only exhibited by their facial expressions. The participants were asked to assess the agents’ personality based on six possible traits.

We found that the participants considered the neutral agent as the most neurotic, while there are also indications that the extraverted agent was considered the most extraverted. We conclude that players will indeed distinguish personality differences between agents based on their facial expression of emotions. Therefore, using a personality model may make it easy for game developers to quickly create a high variety of virtual characters, who exhibit individual behaviors, making them more believable.

I. INTRODUCTION

Over the last decades, computer games have evolved from plain and simple environments, to extensive virtual worlds, with refined visuals and options for players to make personal adjustments [2], [3]. The degree of realism that is achieved by the graphical representation of the game is, however, being shortchanged by the simplistic implementation of the non-player characters (virtual agents) in games, of which the behavior is often controlled by finite state machines [3] rather than responding to the human player. Such agents can hardly be called “believable”. Creating more believable virtual agents is not necessarily about making them as realistic as possible [4]; on the contrary: Bates [5] describes the believability of an agent not as creating a reliable or honest virtual agent, but “one that provides the illusion of life”, so that the player can have a “suspension of disbelief”. Therefore creating a game that is credible is more important than creating one that is realistic [1].

A common solution used by game developers to make virtual agents respond to the human player is to code all possible options in a decision tree [1]. However, creating such a tree, even if only a limited number of options is taken into account, is time-intensive and costly, both during development and testing. Games therefore limit the creation of complex interactions with a few choice non-player characters, turning the remainder into a series of clones which all respond to the player in exactly the same way.

An extensive attempt to tackle the problem of creating believable virtual agents for games was done by Ochs et al. [1], who developed a model which takes the virtual agent’s personality and social relations into account, to determine its emotional state. They claim that their model is consistent with results from (social) psychology. However, they never tested their model in experiments with real life participants, to determine if human players actually consider the agents believable. McRorie et al. [6] evaluated the believability of virtual agents with human participants, but these evaluations were limited to the participants watching pre-recorded footage of interactions with virtual agents, rather than interacting with the agents themselves.

In our research we extend upon the work by Ochs et al. [1] and McRorie et al. [6] by actually testing a model for believable virtual agents in interaction with humans. We use a simplified version of the model of Ochs et al. [1]. We test this model on real life participants, who have to interact directly with the virtual agents, to find out to what extent the agents are indeed believable according to the participants.

II. BELIEVABILITY OF VIRTUAL AGENTS

When discussing the believability of virtual agents, various researchers list different factors of importance. McRorie et al. [6] consider personality as a determinant factor for the believability. Ochs et al. [1] list emotions and social relations. Gebhard [8] add moods to the mix. Regardless the factors, Ortony [7] argues for internal consistency. In practice, the implementation of behavioral characteristics in virtual agents is not based on existing psychological knowledge, but on intuition of the game designers [6]. McRorie et al. [6] reason that an approach to designing virtual agents based on social psychology will result in more believable characters.

In the following subsections, we provide some theoretical background for the most commonly named factors of believability of virtual agents, as well as examples of recent research on modeling these factors.

A. Emotions

The ability to express emotions is key for the believability of virtual agents, as it shows that the virtual agent “cares”
about what is happening in its surrounding environment [5]. Virtual agents who do not show any emotions do not provide an “illusion of life” and therefore do not generate compassion in players.

However, it is not easy to describe what, exactly, emotions are. Despite being extensively investigated, researchers have not yet agreed on a uniform definition of emotions. There is no agreed-upon clear distinction between the terms “emotion”, “mood”, “attitude” and “feeling”[9]. Literature often focuses on Ekman’s notion of six basic emotions: happiness, sadness, fear, anger, surprise and disgust [10]. Du, Tao and Martinez [11] describe a group of compound emotions, which can be constructed by combining basic emotions, such as “happily surprised” and “sadly disgusted”.

In designing virtual agents with believable emotions three main challenges occur: (1) identifying the specific circumstances in which a certain emotion should appear, (2) determining how to display these emotions, and (3) defining which behavior on the part of the agent will be the result [1].

As for identifying circumstances, the Ortony, Clore and Collins (OCC) model provides a popular approach to appraise events in research concerning the simulation of emotions [12]. The OCC model determines which of 22 types of emotion occurs using three different factors: the consequences of events, the actions of the agents, and the aspects of objects [13].

Displaying emotions in virtual agents is usually attempted via facial expressions, gestures, and tone of voice, with facial expressions being most important [14]. A well-known system that can be used for animating facial expressions in virtual agents is the Facial Action Coding System (FACS) [15]. The system measures facial expressions in terms of Action Units (AUs), which are the smallest possible units of moving muscles in the face. For each facial expression, certain combinations of AUs are activated with different intensities.

Amini et al. [16] developed the HapFACS software, which can simulate AUs described by FACS on virtual agents. The system works together with Haptek software, which creates virtual avatars. The advantage of using the software is that the AUs are FACS validated [16]. The main disadvantage is that it only runs on outdated operating systems.

B. Personality

Personality can be described as “a pattern of behavioral, temperamental, emotional, and mental traits for an individual”, including its “longterm tendencies” [13]. It can be interpreted by “encoding and decoding [...] mainly non-verbal cues” [14]. According to Albeck and Badler [13], personality is key to creating believable autonomous agents. Ochs et al. [1] state that virtual agents with a personality will contribute to the consistency of a game, and thus to believability of the agents.

There are two well-known psychological models that describe personalities. The first is the model by Eysenck [17], which uses only two traits: extraversion versus introversion and neuroticism versus emotional stability. The second (and most commonly used) model is the five-factor model by Costa and McRae [18], which adds to Eysenck’s model the factors openness, agreeableness, and conscientiousness.

McRorie et al. [6] explain how people tend to make automatic judgments about the personalities of their interaction partners, so they should also do this with virtual agents. Modern game characters already feature facial expressions and head and eye movements, on which players will base personality judgments. To test their theory, McRorie et al. [6] used still images of their designed virtual agents with different personalities as well as video clips, where the interaction between the virtual agents with a (male) player was shown. They measured three values: the virtual agents’ believability, consistency and familiarity. However, none of the participants in the study interacted with the virtual agents themselves.

C. Moods

Moods are often considered similar to, but not the same as, emotions. Ekman [10] points out several differences between moods and emotions. The three main differences are: (1) moods last longer than emotions (though it is unclear what lengths of time are meant), (2) moods do not have a unique facial expression, contrary to emotions, and (3) moods may have an influence on emotions. As an example for the third difference, consider irritation (a mood) and anger (an emotion). When one is irritated, one is quicker to get angry, and being irritated causes the anger to last longer.

Gebhard [8] used mood as a factor in his “layered model of affect”, called ALMA, which was intended to simulate believable behavior in virtual agents. According to Gebhard, the three types of affect are emotions (short-term affect), moods (medium-term affect), and personality (long-term affect). These three types interact with each other in the ALMA model, which was implemented in virtual agents. However, whether or not this model improves believability of agents has not been investigated yet.

D. Social relations

Most models of social relations contain a finite set of variables, each of them representing one of the dimensions of the social relation between two (virtual) agents [1]. Ochs et al. [1] state that each model uses different variables; there are no pre-set requirements that researchers have reached consensus on. An example of the implementation of social relations in a game is found in the work of McCoy et al. [19].

E. Ochs’ model

Ochs et al. [1] did an extensive attempt to design virtual agents based on knowledge of the field of (social) psychology. Instead of having a focus on a single component that is important for the believability of a virtual agent (emotions, personality, moods or social relations), they developed a model, aimed at game developers, which calculates the virtual agents’ emotional state and its resulting behavior, by taking the virtual agents’ personality and social relations into account. They argue that this results in more consistent behavior, increasing the believability of the virtual agents.
Ochs’ model (see Figure 1) is based on several theories in (social) psychology. It uses an adapted version of the OCC model to model emotions. Two differences with the OCC model is that Ochs’ model uses 10 instead of 22 emotions (namely joy, hope, disappointment, distress, fear, relief, pride, admiration, shame, and anger), and that it does not require the definition of explicit agent goals (because, as Ochs reasons, they might often be difficult to apply for a game). Instead of goals, Ochs’ model uses “actions, objects, and other characters of the environment” [1].

For personalities, Ochs’ model uses the original Eysenck two-factor model. The influence of personality on emotions is based on Gebhard’s ALMA model [8] (despite the fact that ALMA is used with moods rather than emotions). In Ochs’ model, the intensity of an emotion is determined by the personality dimensions, as follows: a higher score on the extraversion dimension increases the intensity of “positive” emotions (such as joy, hope, pride, and relief), while a higher score on neuroticism increases the intensity of “negative” emotions (such as sadness, frustration, irritation, and anger). This idea of linking the effect of personality to emotions in this way follows from the work by Watson and Clark [20].

Social relationships are covered by Ochs’ model by calculating the effects of “liking”, “dominance”, “solidarity”, and “familiarity”.

While Ochs’ model is quite extensive, we note three shortcomings in its design and the reasoning behind the design:

1) No motivation is given for reducing the number of emotions from 22 to 10, and for adapting the ALMA model from moods to emotions;
2) While it is claimed that the model will result in believable, consistent behavior, the model omits a description of how behavior follows from the model; and
3) The model was never tested with real-life participants; the only tests done with the model were on correctness of implementation, not on the effect on the believability of agents controlled by the model.

In our research, we aim to (partly) resolve the third of these shortcomings, by testing the effect of agents controlled by a version of Ochs’ model on humans interacting with these agents.

III. EXPERIMENTAL SETUP

For our experiments, we derived a simplified version of Ochs’ model [1] to control agents with a predefined personality in an interaction with human participants.

A. Model design

The primary goal of our research is to investigate if personalities in a virtual agent, simulated by our model, are recognized by real-life participants. Since we focus on the recognition of the virtual agents personality, we have to exclude other interfering aspects of the model of Ochs et al. [1]. The simplified version of Ochs’ model that we used is shown in Figure 2. There are two important differences between our model and Ochs’ model:

1) In our model we do not account for attitudes. In Ochs’ model, attitudes influence the intensity of emotions. Since we want to measure the effect of personality, which also influences the intensity of emotions, we have to leave attitudes out of the model.
2) Neither do we account for social relations in our model. In Ochs’ model, social relations influence behavior next to emotional states. Since in our model we want to measure the effect of personality, which interacts with emotions which influence behavior, we need to exclude any other factors that influence behavior.

As can be seen in Figure 2, in our model the virtual agent is characterized by its personality, which influences the intensity of emotions triggered by events, and thus the emotional state of the virtual agent. The agent’s emotional state determines its expression of these emotions.

Emotions can be expressed in different ways, such as visual appearance and behavior; however, in our experiments
we focus exclusively on facial expressions. Moreover, rather than including all ten emotions covered by Ochs’ model, we chose to focus on only two combinations of emotions: joy-distress and admiration-anger. The reason to select these two combinations is that they are relatively easy to trigger and recognize, as they are direct reactions to events occurring at a present moment, instead of being related to a larger context of the past or future (according to the definition of these emotions by Ochs et al. [1]).

B. Implementation

We implemented our model in a small game, in which the player has an interaction with three different virtual agents. The model allows the creation of a wide variety of agents, but for our experiments we decided to focus on three which are fundamentally different: (1) an extraverted agent which scores high on the extraversion dimension, but neutral on the neuroticism dimension; (2) a neurotic agent which scores high on the neuroticism dimension, but neutral on the extraversion dimension; and (3) a neutral agent, which scores neutral on both dimensions.

Each agent has an emotional state, consisting of values for joy, distress, admiration, and anger. In our game, the values for the emotions in the emotional state range from 0 to 2, with an initial value of 0.

Events in the game are defined as 3-uplets consisting of (1) an independent event that occurs or an agent that performs an action, (2) a positive or negative effect that follows, and (3) the virtual agent that experiences the effect. This information is used to determine which emotion occurs following the event. Emotions that may follow an independent event are joy (positive) or distress (negative). Emotions that may follow the action of an agent are admiration (positive) or anger (negative).

After the type of emotion is determined, the intensity is calculated. In our game the initial intensity of each emotion is 1, which is further influenced by the personality of the agent. For the extraverted agent, for positive emotions the intensity value is doubled, while for the neurotic agent, for negative emotions the intensity is doubled. For agents of other personality types (e.g., slightly neurotic but also a bit extraverted) obvious adaptations to these values can be made.

The experienced emotion’s intensity is then added to the corresponding value in the agent’s emotional state, in our implementation capped at the boundary of 2.

After every event in the game, all values in the emotional state of the agent are decreased by a “decrease rate”. In Ochs’ model, a logarithmic decrease rate is used, which we simplified to a linear rate (of which the effect happened to be close to what happens in Ochs’ model), which decreases each of the values by 0.2 with a lower bound of 0.

Finally, the agent’s emotional state determines its facial expression in the game. For this, we took the dominant emotion (i.e., the one with the highest value) from the emotional state, and selected a facial expression corresponding to the intensity of that emotion. Examples of facial expressions are shown in Figure 3: rows from top to bottom representing joy, distress, admiration, and anger, columns from left to right representing intensity values 0.4, 0.8, 1.2, 1.6, and 2.0. The emotional expressions were generated with the FaceGen software.

FaceGen has pre-programmed expressions for “anger”. For “distress”, we used the pre-programmed “sadness” expressions. For “joy”, according to Du et al. [11] AUs 6, 12, and 25 are used, which we managed to approach using detailed FaceGen sliders. For “admiration” we consulted Parke and Waters [21], which describe an admiration expression as generated by “the muscles of astonishment associated with those of joy”, which we also implemented using detailed FaceGen sliders. We pre-tested the expressions with ten participants, asking them to describe the emotion associated with a facial expression. Most participants labeled our initial creations correctly, except for “admiration”. We tweaked the expression for “admiration” and offered it to a second group of ten test participants, of which five specifically recognized the face as expressing “admiration”, and none of the others labeled it as a negative emotion. This was the expression that we used in the experiments.

Note that we are not claiming that personality and emotions only influence facial expressions in practice: in fact, they influence all aspects of interaction, including speech patterns [22], gestures [23], and behavioral responses [18]. However, to be able to correctly assign the players’ assessments to particular elements in the game, we wanted to limit the effects of the model to a single feature, for which we chose facial expressions.

C. Game

We implemented our game in Ren’Py, a tool for creating visual novels. Players control the game by making menu choices in their interaction with virtual agents. The simple storyline of our game had the participant fulfill the role of a detective who is investigating a robbery. The player had to interrogate three virtual agents: the extraverted, neurotic, and neutral agents that we defined. Two short excerpts from a dialogue are displayed in Table I, including the emotional effect of the player’s statements on the agent, and whether this effect was caused by an event, or an action of the player. These excerpts provide only one possible choice for each of the player’s dialogue options. Note that in the actual game the dialogue is much longer, and includes a situation in which the player suspects the agent of being involved in the robbery.

The choices that the player made had little effect on the progression of the story; while it might seem to the participant that their choices drove the interrogation, they actually always led to the same follow-up responses of the virtual agents. The interactions were slightly different between the three agents that the player encountered, but the order in which the interactions were presented to the participants were always the same. However, the three personalities imprinted on the agents differed between the participants: the six different orders in which the three personalities could be presented to the participants all occurred the same number of times. For example, while the first agent encountered always had
a particular interaction with the participant, one-third of the participants had this interaction with the extraverted agent, one-third with the neurotic agent, and one-third with the neutral agent. The emotional impact of the events was only visible in the facial expressions of the agents.

In each interaction, five different events occurred that triggered an emotion. Three of those were negative, one of them caused by the participant. Two were positive, one of them caused by the participant.

After each of the three interactions with the virtual agents, the participants had to indicate to what extent they thought the virtual agents had certain personality traits. This was questioned by means of five-point Likert scales, where one resembled “totally disagree” and five resembled “totally agree”. The participants did this for the following six statements: (1) the agent is introverted, (2) the agent is emotionally stable, (3) the agent is orderly, (4) the agent is extraverted, (5) the agent is neurotic, and (6) the agent is intelligent. The third and sixth statement served as control statements: in theory, the scores on these two statements should not differ between the personalities.

At the end of the game, the participants had to answer a few questions on their gender, age, education and experience with games.

D. Participants

36 people participated in the experiment, who were between 21 and 57 years old ($M = 27.31, SD = 10.11$). Of these participants, 16 were male and 20 female. Four people attended higher professional education, 32 went to university. 16 of the participants stated they did not have any experience with computer games at all. The participants were equally distributed over the six conditions.
and “neurotic” was reversed. The internal consistency of the traits were combined into six general (dependent) variables. A principal factor analysis was conducted, resulting in three factors. Factor loadings are displayed in Table II. The clustered items indicated that the first factor represents the introverted-extraverted trait, the second factor represents the emotionally stable-neurotic trait, and the third factor represents the two control variables: orderly and intelligent. We added these factors as three new variables, consisting of (1) the mean score of extraverted and reversed introverted, (2) the mean score of emotionally stable and reversed neurotic, and (3) the mean score of orderly and intelligent.

For the nine variables (the six original scores and the three new ones) we did an analysis of variance with repeated measures. Table III shows the means and (between parentheses) standard deviations for each of the scores.

The analysis of variance shows the following:

- Between the agents, there is no significant difference for the scores on “introverted”, “orderly”, “extraverted”, “intelligent”, “introverted-extraverted”, or “orderly-intelligent”.
- Between the agents, there is a significant difference for “neurotic” ($F(2, 70) = 7.02, p < .005$). The score for the neurotic agent is significantly higher than the scores for the other two agents. There is no significant difference between the scores for the extraverted and neutral agents.
- Between the agents, there is a significant difference for “emotionally stable” ($F(2, 70) = 8.72, p < .001$). The score for the neurotic agent is significantly lower than the scores for the other two agents. There is no significant difference between the scores for the extraverted and neutral agents.

### IV. RESULTS

Figure 4 shows the average assessment of the participants over the personality traits of the three virtual agents (exact numbers follow in Table III, and more numerical details have been reported by Kersjes [24]). The six personality traits that are assessed are shown in groups of three bars, of which the left represents the score for the extraverted agent, the middle for the neutral agent, and the right for the neutral agent. We stress once more that by balancing the order in which the virtual agent personalities were presented to the participants, the differences between the assessments are the sole result of the participants’ observations of the facial expressions of the virtual agents, which were generated by our personality model.

A visual inspection shows some notable differences between the assessments:

- All three agents score higher on “extraverted” than on “introverted”. This indicates that all of them are considered to be leaning towards being extraverted, though for the extraverted agent the difference between the scores is a bit higher than for the other two agents.
- Both the extraverted and neutral agent score higher on “emotionally stable” than on “neurotic”. For the neurotic agent, however, these scores are almost equal. This indicates that the participants assess the neurotic agent as considerably more neurotic than the other two agents.
- Finally, the extraverted agent scores higher on both “orderly” and “intelligent” than the other two agents. The neurotic agent scores lowest on these traits, though the difference with the neutral agent is small.

### A. Reliability and factor analysis

In order to further analyze these results, a reliability and factor analysis was conducted. All scores on each of the six traits were combined into six general (dependent) variables. The personality of the agents served as an independent variable in this way of grouping the data. The polarity of “introverted” and “neurotic” was reversed. The internal consistency of the six measured traits was low (Cronbach’s $\alpha = .45$), thus all items appeared worthy of retention.

A principal factor analysis was conducted, resulting in three factors. Factor loadings are displayed in Table II. The clustered items indicated that the first factor represents the introverted-extraverted trait, the second factor represents the emotionally stable-neurotic trait, and the third factor represents the two control variables: orderly and intelligent. We added these factors as three new variables, consisting of (1) the mean score of extraverted and reversed introverted, (2) the mean score of emotionally stable and reversed neurotic, and (3) the mean score of orderly and intelligent.

For the nine variables (the six original scores and the three new ones) we did an analysis of variance with repeated measures. Table III shows the means and (between parentheses) standard deviations for each of the scores.

<table>
<thead>
<tr>
<th>Scale</th>
<th>Introverted</th>
<th>Emotionally stable</th>
<th>Orderly</th>
<th>Extraverted</th>
<th>Neurotic (reversed)</th>
<th>Intelligent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introverted (reversed)</td>
<td>.83</td>
<td>.73</td>
<td>.24</td>
<td>.11</td>
<td>.77</td>
<td>.12</td>
</tr>
<tr>
<td>Emotionally Stable</td>
<td>.73</td>
<td>.54</td>
<td></td>
<td>.44</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Orderly</td>
<td>.11</td>
<td>.34</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Extraverted</td>
<td>.77</td>
<td>.84</td>
<td>.10</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Neurotic (reversed)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intelligent</td>
<td>.12</td>
<td></td>
<td>.67</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---

**TABLE I**

**TWO EXCERPTS FROM A DIALOGUE.**

<table>
<thead>
<tr>
<th>Speaker</th>
<th>Text</th>
<th>Effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Player:</td>
<td>“You have a personal storage locker at the bank. Is that correct?”</td>
<td></td>
</tr>
<tr>
<td>Agent:</td>
<td>“Yes.”</td>
<td></td>
</tr>
<tr>
<td>Player:</td>
<td>“Can you tell me what was in it?”</td>
<td></td>
</tr>
<tr>
<td>Agent:</td>
<td>“Why? What is this about?”</td>
<td></td>
</tr>
<tr>
<td>Player:</td>
<td>“The bank was robbed this weekend. I am investigating the robbery.”</td>
<td></td>
</tr>
<tr>
<td>Agent:</td>
<td>“I understand. There were a few family valuables in the locker. Jewelry and a watch.”</td>
<td></td>
</tr>
<tr>
<td>Player:</td>
<td>“Some of the personal lockers were broken into. I am afraid that yours was also emptied.”</td>
<td></td>
</tr>
<tr>
<td>Agent:</td>
<td>“Will I ever see my possessions again?”</td>
<td></td>
</tr>
<tr>
<td>Player:</td>
<td>“I can tell you that some of the stolen goods were recovered this morning.”</td>
<td></td>
</tr>
<tr>
<td>Agent:</td>
<td>“Did you also find my jewelry?”</td>
<td></td>
</tr>
<tr>
<td>Player:</td>
<td>“If you give me a moment, I will check.”</td>
<td></td>
</tr>
<tr>
<td>Player:</td>
<td>“My colleague tells me that it looks like some of your possessions have indeed been found. I will make sure that they are returned to you as soon as possible.”</td>
<td></td>
</tr>
</tbody>
</table>

---

**TABLE II**

**FACTOR ANALYSIS RESULTS.**

---

**FIG. 4.** Assessment of personality types.
assessed the second agent as significantly more introverted for being extraverted were less pronounced than the results for other two agents. We offer two possible reasons why the results indicate that this personality scored higher on “extraverted” and lower on “introverted” than the other two agents. There is no significant difference between the scores for the extraverted and neutral agents.

B. Conversation order

While the personalities were presented in varying orders to the participants, the textual aspects of the story always occurred in the same order. To see whether there was a difference between the three conversations that the participants had, we did an analysis of variance for these conversations. We found a significant difference for the “introverted” scores $(F(2, 70) = 10.15, p < .001)$. The score for the neurotic agent is significantly lower than the scores for the other two agents. There is no significant difference between the scores for the extraverted and neutral agents.

### Table III

<table>
<thead>
<tr>
<th>Trait</th>
<th>Extraverted</th>
<th>Neurotic</th>
<th>Neutral</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introverted</td>
<td>2.50 (.97)</td>
<td>2.53 (.88)</td>
<td>2.64 (.90)</td>
</tr>
<tr>
<td>Emotionally stable</td>
<td>3.56 (.70)</td>
<td>2.86 (.93)</td>
<td>3.31 (.75)</td>
</tr>
<tr>
<td>Orderly</td>
<td>3.36 (.87)</td>
<td>3.14 (.80)</td>
<td>3.17 (.74)</td>
</tr>
<tr>
<td>Extraverted</td>
<td>3.31 (.89)</td>
<td>3.19 (.82)</td>
<td>3.14 (.93)</td>
</tr>
<tr>
<td>Neurotic</td>
<td>2.31 (.67)</td>
<td>2.89 (1.01)</td>
<td>2.33 (.72)</td>
</tr>
<tr>
<td>Intelligent</td>
<td>3.42 (.65)</td>
<td>3.22 (.54)</td>
<td>3.33 (.76)</td>
</tr>
<tr>
<td>Introverted-extraverted</td>
<td>3.40 (.85)</td>
<td>3.33 (.74)</td>
<td>3.25 (.85)</td>
</tr>
<tr>
<td>Emotional stable-neurotic</td>
<td>3.63 (.58)</td>
<td>2.99 (.90)</td>
<td>3.49 (.65)</td>
</tr>
<tr>
<td>Orderly-intelligent</td>
<td>3.39 (.62)</td>
<td>3.18 (.56)</td>
<td>3.25 (.58)</td>
</tr>
</tbody>
</table>

V. Discussion

We found significant differences between the assessment of the participants of the agents’ emotional stability and neuroticism. The agent for whom, based on our model, the facial expressions were created representing a neurotic personality, was deemed, by the participants, as considerably more neurotic than the other two agents. Note that the main effect of having the neurotic personality was that the agent showed a more intense response to negative events than the other two agents, while showing the same response to positive events as the neutral agent. We may therefore conclude that the model manages to display a recognizable neurotic personality.

The participants did not clearly recognize the extraverted personality, though the individual scores that the extraverted agent received seem to indicate that this personality scored higher on “extraverted” and lower on “introverted” than the other two agents. We offer two possible reasons why the results for being extraverted were less pronounced than the results for being neurotic.

Firstly, as explained above, we found that the participants considered the second agent they encountered as significantly more introverted than the other two agents. This indicates that the textual cues of the second conversation were representative of an introverted personality. While we had intended the results to only depend on facial expressions, we now know that the results for introversion and extraversion were influenced by the conversation texts. In future research, this can be resolved by not only varying the order of the agent personalities, but also the order of the conversations.

Secondly, each conversation contained three events that triggered negative emotions (which make a difference for the neurotic agent), while only two events triggered positive emotions (which make a difference for the extraverted agent), which also occurred late in the conversations. In retrospect, this might have caused the neurotic traits to be recognized more easily. In future research a longer and more diverse interaction with the agents may make it easier to recognize all personality traits.

One interesting observation on our results is that, while we included two control traits that should not differ between the personalities, from Figure 4 we can see that the participants considered the extraverted agent to be more orderly and more intelligent than the other two agents, and that the neurotic agent was considered less orderly and less intelligent than the other two. Evidently, subconsciously people automatically associate extraversion with being orderly and intelligent.

As for the use of a personality model by game developers: the idea is that the model determines emotional state based on settings for personality (and perhaps also moods and social relations). This emotional state can then be used to determine not only facial expressions, but also actual behaviors, body language, tone of voice, and choice of words. In our research we only used facial expressions to allow us to draw solid conclusions on participants recognizing personality features; however, game developers naturally would want to include more extensive influences of emotional states. This would help the diversity and believability of the portrayed characters.

VI. Conclusion

One approach to create more believable characters in games is to provide characters with a personality, which automatically influences the characters’ behavior and (emotional) expressions without extra coding on the part of the game developers. For instance, two guards could have exactly the same programmed conversations with the player, but by using a personality model to generate their visual characteristics (such as facial expressions), these two guards can give the impression to the player of being two fundamentally distinct characters. In a more extensive implementation where also behaviors are concerned, one guard could have a personality that responds more intensively to negative emotions than the other, and thus be quicker to get angry and try to arrest the player.

Few personality models for game characters are known, and as far as we know none of them have been tested.
with human participants interacting with virtual characters. Therefore, there is little evidence that humans indeed think the virtual agents driven by such a personality model behave in a believable or realistic way. In our research, we presented a simplified version of the model of Ochs et al. [1], which we use to let three different agents interact with human participants in a game; the three agents have an extraverted, neurotic, and neutral personality. In our experimental setup, personality only influences the facial expressions of the agents.

Our results showed that the neurotic agent was indeed deemed to be more neurotic than the other two agents. While the results for the extraverted agent were not significant, a visual inspection (see Figure 4) shows that it was deemed to be more extraverted than the other two agents. It should be noted that the results for the extraverted agent might have been influenced by the fact that the second agent that the participants interacted with was considered to be more introverted than the other two, regardless of its actual personality. This issue may be resolved in future research.

We conclude that using a personality model to control the behavior and emotional expressions of virtual characters may indeed have human players regard these characters as fundamentally different with recognizable personalities, even if the interaction or storyline remains unchanged. Therefore, using a personality model may make virtual characters more diverse, more human-like, and thus more believable.

REFERENCES