A Process Model of Empathy For Virtual Agents

Abstract

For more than a century, empathy has been a central topic in the study of human emotion. It plays a crucial role in our everyday social life, having implications for the survival of the species. In the case of agents that inhabit virtual worlds and interact socially among each other and with humans, empathy has also been considered to be an important mechanism to promote engaging and believable interactions. However, creating empathic agents, until recently, has been accomplished mostly through the implementation of specific empathic behaviors or by using domain-dependent empirical models. In this article, we propose a generic computational model of empathy that is grounded on recent psychological theories about empathy. The proposed model treats empathy as a process in which the intensity of the empathic response is modulated by a set of factors that involve the relationship between the agents of the empathic interaction, namely, the similarity and affective link, as well as some characteristics of the empathizer agent, such as mood and personality. This model was implemented into an affective agent architecture, which was then used in an evaluation that had 77 participants. The results indicate that our empathy model, when used to simulate a social scenario with a small group of agents, significantly changed the way that the users perceived and described the interactions between those agents.

Keywords: Empathy, Process Model, Virtual Agents, Empathic Behavior

1. Introduction

In recent decades, we have witnessed a significant increase in computational models of emotions for autonomous agents. From the seminal work of the OZ group (Bates et al., 1994), almost twenty years ago, to the recent developments of the past few years, we have witnessed a "seemingly bewildering array of complementary computational models of emotion (Marsella et al., 2010)". The ultimate aims of these models are not only to provide useful insights into emotion research but also to facilitate the social interaction between humans and machines. However, to create these computational models of emotion, we must consider the interplay between many processes that are associated with emotions, such as perception, appraisal, imitation, learning, and, most certainly, empathy.

Broadly speaking, "empathy" relates to the capability of perceiving, understanding and experiencing another person's emotions. Throughout history, scientists have extensively investigated this ability and its underlying processes. Consequently, many definitions of empathy have been proposed, with each definition incorporating diverse concepts and emphasizing different perspectives. However, although there is no existing consensual definition, the importance of empathy is indisputable because it plays a vital role in human social behavior as well as in the behavior of some non-human species. From an evolutionary point of view, empathy acts as a critical factor to promote survival (Plutchik, 1987). From a social point of view, empathy is strongly associated to altruism, moral judgement, prosocial behavior, and cooperation (Hoffman, 2001, 1987; Batson and Shaw, 1991). As pointed out by Pizarro et al. "without empathy, more complex moral emotions such as guilt and anger on behalf of others would most likely not exist" (Pizarro et al., 2006). In summary, empathy has long been considered to be a major contributor to successful human social interaction.

Given the importance of empathy in humans, many researchers in the fields of virtual agents and social robots have given it significant attention. The use of agents with the ability to express empathy toward the user is continuously growing, specifically in applications for education, games and counseling (Bickmore, 2003; Prendinger and Ishizuka, 2005; Paiva et al., 2005; Ochs et al., 2009; Deshmukh et al., 2013). A study conducted by Brave et al. (Brave et al., 2005) showed that both female and male participants who had a short-term interaction with an empathic agent significantly perceived the agent as more caring, likeable, and trustworthy than those who interacted with a non-empathic version of the same agent. In the area of robotics, empathy has also recently been addressed. Researchers are studying how robots can trigger empathic reactions from users as well as giving robots mechanisms for displaying empathic behaviors (Kozima et al., 2004; Tapus et al., 2007; Leite et al., 2012).

However, in spite of the growing interest in modeling empathy in virtual agents, thus far the main focus has been on studying users' reactions' to the display of specific empathic behaviors during a given task. The agent's empathy is usually triggered by a set of domain specific rules, sometimes authored in an ad-hoc manner. The main disadvantage of using this type of approach is that it offers little flexibility and significantly narrows the scope of possible empathic interactions.

In this paper, we present a computational model of empathy that is inspired by neuropsychological theory (de Vignemont and Singer, 2006). The main goals are to facilitate the emergence of empathy in virtual agents and to eventually help to better understand its process in the interaction between humans and machines. From an Artificial Intelligence perspective, the proposed model aims to enhance affective agent architectures based on appraisal theories, (e.g FAtiMA (Dias and Paiva, 2005b), EMA (Marsella and Gratch, 2009)), by adding an explicit process of empathy that automatically elicits and modulates empathic emotions based on the appraisal mechanisms that already exist. Moreover, the model is defined in a general way; it enables virtual agents to empathize with any other agent, thus not only with users but also amongst themselves. In fact, in current multi-agent virtual environments, empathy between agents is often left unexplored. We believe that a process model of empathy can be used to enrich multi-agent virtual environments, which would allow a broader set of interaction experiences for all involved and would, thus enable a stronger user engagement.

An implementation of the proposed empathy model was accomplished by extending FAtiMA(Dias and Paiva, 2005b), an existing emotional agent architecture, which is grounded on the OCC appraisal theory of emotions (Ortony et al., 1988). To implement empathy in this architecture, an additional appraisal process was added, which is appropriately named *Empathic Appraisal*. Using this implementation, an evaluation was conducted to measure the psychological effect of the generated empathic behavior on users. This study is based on a small multi-agent scenario in which the model was applied to produce empathic interactions between the virtual agents. The scenario is composed of a short emergent story, which is enacted by a small group of autonomous agents. It reuses the same virtual environment and technological framework used by the FearNot! educational application (Paiva et al., 2005).

Given the nature of the subject that is being modeled and for the purpose of testing the impact that the architecture had on generating appropriate believable behavior, we conducted an evaluation that focused on the perception of the empathic behavior that was generated by the architecture.

This article is organized as follows. In Section 2, we describe important findings in empathy theory that inspired and grounded the approach provided here. Then, the proposed model is presented as well as its integration into an agent architecture. In Section 5, a scenario used for exploring the empathic architecture is described. In the subsequent section, the experimental results of two different evaluations of the model are presented and analyzed. We then place this current work into the context of a wider picture of research on emotion architectures and empathy by addressing important related work. Finally, we discuss the limitations of the work presented, draw some conclusions and envisage future work.

2. Background

Empathy is pervasive in our society. We can see newspapers making headlines out of the possible existence of empathy in rats, the lack of empathy in current companies, or the importance of having the right empathic edge for winning or losing presidential elections. Hoffman defines the concept as "an affective response more appropriate to another's situation than one's own" (Hoffman, 2001). Similarly, Stotland defines empathy as "an observer's reacting emotionally because he perceives that another is experiencing or is about to experience an emotion"(Stotland, 1969). Although there are many more definitions of empathy, the ones just mentioned focus on the notion that empathy involves experiencing emotional responses that are more congruent with someone else's situation. For example, when we hear a story about an elderly couple that lost their pensions and lifetime savings in the recent Cypriot economic crisis, we might feel sad and upset, even though the story has nothing to do with us. However, we could feel substantially more upset and most likely quite angry at bankers and politicians if that couple is from our family or if what happened to them is likely to happen to us. Empathy is at the center of these emotional responses.

The well known "Perception-Action Hypothesis" (Preston and de Waal, 2002) states that the perception of a behavior in another person (the target) will automatically activate one's own representation for that behavior. In other words, the observer's perception of the other's emotional state is linked with his or her own somatic and autonomic responses via his own neurological representations. This general perception-action model of empathy goes even farther by considering that even the imagination of someone else, in a specific emotional state, will automatically activate a representation of that state in the observer (empathic appraisal), leading to an empathic activation and response (for example, imagining the suffering of the elderly Cypriot couple). Adopting this view of empathy as a high-level process allows that some empathy related phenomena such as emotional contagion[1], prosocial behavior[2], sympathy[3], and perspective taking[4], can share the same base structural model that relies on the perception-action mechanism (Preston and de Waal, 2002). Another researcher who supports the view that empathy should be regarded as a combination of different inter-related processes is Davis (Davis, 1994). De Waal (de Waal, 2007) uses the Russian doll as a metaphor to exemplify the concept of empathy as a set of nested layers that subsume not only simple mechanisms at its core (such as mimicry) to more complex mechanisms, cognitive filters, and perspective-taking at its outer layers.

However, as rationalized in (de Vignemont and Singer, 2006), it is possible to group some general tendencies to characterize empathy around two main views: a broader, more inclusive view, in which where empathy is seen more like an umbrella for many related phenomena spanning from emotional contagion through cognitive perspective-taking, and a narrower view that distinguishes empathy from the other related concepts, using a more precise and clear definition. This narrower view of empathy comes from the work of de Vignemont and Singer (de Vignemont and Singer, 2006), which considers that empathy exists only if: (i) the empathizer is in an affective state; (ii) this state is isomorphic to the target's affective state; (iii) this state is elicited by the observation or imagination of the target's affective state; and (iv)

[1]"Emotional contagion" is defined by (Rapson et al., 1993) as the tendency to automatically mimic and synchronize facial expressions, vocalizations, postures and movements with another person, and consequently, converge emotionally. Hoffman (Hoffman, 2001) considers that emotional contagion is type of "immature" empathy.

[2]Prosocial and altruistic behavior is a result of our genuine concern for the well-being of the others (Batson and Shaw, 1991)

[3]According to N. Eisenberg, sympathy is "an emotional response stemming from another's emotional state or condition that is not identical to the other's emotion, but consists of feelings of sorrow or concern for another" (Eisenberg and Miller, 1987). According to Baron-Cohen et. al. sympathy is an instance of the affective component of empathy and occurs when the observer's emotional response to the distress of another motivates that observer to take an action to alleviate the other person's suffering (Baron-Cohen and Wheelwright, 2004).

[4]Perspective taking (see (Hoffman, 2001)) can be seen as one of the empathic appraisal modes that arises when the observer imagines him or herself in the situation of the target and imagines how he or she would feel in that same situation.

the empathizer knows that the target is the source of his own affective state. This view leaves out common related concepts of empathy, such as sympathy, emotional contagion, personal distress and cognitive perspective-taking.

In spite of these different ways of classifying what is and what is not empathy, most researchers agree that empathy is not the same for everyone, nor is it the same for different situations and contexts. Many factors mediate the arousal of empathy. De Vignemont and Singer (de Vignemont and Singer, 2006) have identified several different factors that modulate an empathic response and have grouped them into the following four categories:

1. **Features of emotions** - The type and features of the emotion of the target affect the empathic process and response. For example, it was shown in (Saarela et al., 2007) that the amplitude of the empathic responses by observers is modulated by the intensity of the target's facially displayed pain that is related to the emotion. Additionally, according to (de Vignemont and Singer, 2006) , we empathize more with primary emotions such as fear, happiness or sadness than with secondary emotions such as jealousy.

2. **Relationship between empathizer and target** - Aspects such as how familiar or similar a person is or the affective link that there is between the two leads to different empathic responses from the observer. It is well known that there are ingroup-outgroup biases even in involuntary responses, such as empathy for pain, or yawn contagion. In fact, this bias even extends to some of our close non-human primates (chimpanzees), as shown by F. de Wall (Campbell and de Waal, 2011).

3. **Situative context** - The situation in which the events occur affects tremendously the response of the observer. For instance, an observer's empathic response can be influenced by the presence of others in the immediate situation. Many diverse studies on this well known “bystander effect” have demonstrated in a consistent manner that the presence of others inhibits helping behaviors in the observers. In fact, according to (Latane et al., 1968) as the number of people that are present in a situation increases, the less compelled or responsible the observer is to help.

4. **Characteristics of the empathizer** - Individual differences of the empathizer affect the empathic response. More specifically, the mood,

personality, gender, age, emotional repertoire, and emotional regulation capacities are among some of the individual characteristics that could affect the response. For example, it has been shown that people who score high on social responsibility also score high on dispositional empathy.

Our process model of empathy, although borrowing the idea and structure proposed by the Perception-Action Hypothesis" (Preston and de Waal, 2002), takes a small step in creating empathic agents by adopting some elements of the definition proposed by de Vignemont and Singer and implementing some of the mediation factors to modulate the empathic responses between virtual agents.

3. Computational Model of Empathy

For the purpose of a developing a computational model of empathy for autonomous agents, we consider that there are two main issues the model should address. The first is characterizing the empathic response and the second is describing the nature of the inner mechanisms of the empathic process through which the response is generated. Our proposed model addresses each of these issues in their own separate phase: the *Empathic Appraisal* and the *Empathic Response* (see Figure 1). The former is responsible for the elicitation and modulation of an *Empathic Emotion*, while the latter is responsible for generating a possible *Empathic Action*. However, before these two processes are described, we first discuss the notions of appraisal, emotions and events, which are fundamental for the proposed model.

The concept of evaluating an event in relation to one's goals and desires goes back to almost 2,500 years ago, with the ideas put forward by Aristotle. However, it was in the fifties that Magda Arnold proposed the notion of "appraisal" as judgements made by an individual about how good or bad an event is (Arnold and Gasson, 1954). This appraisal is associated with the individual's goals and motivations, and as such, depends on the way that a person relates to the environment. Since then, distinct appraisal theories have been proposed with different concepts about the structure and/or the process of appraisal (an extensive review on several theories is given in (Roseman and Smith, 2001)). A dividing issue among these appraisal theories is whether emotions can be divided into discrete categorical types or if they have a continuous nature instead. For the purpose of our empathy model,

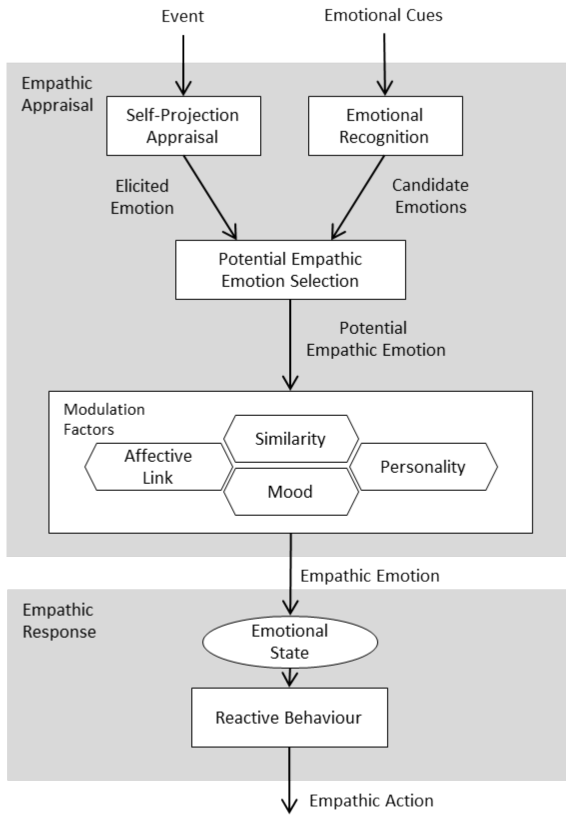

Figure 1: Empathy Model Diagram

the concept of emotion chosen falls in the first category. More precisely, emotion is formally defined as a tuple $< type, valence, intensity, cause >$ where: *type* is the name of the emotion type (e.g., Joy, Distress, Anger); *valence* indicates whether the emotion has a positive or negative valence (1 or -1), which is directly associated with its type (e.g., Joy is positive, Anger is negative); *intensity* is a number greater than zero that corresponds to the emotion's intensity; and *cause* corresponds to the event that caused the emotion.

With respect to the events that occur in the virtual environment, they are formally defined as a tuple $< subject, action, target, parameters >$, where: *subject* is the name of the agent that performed the event's action; *action* is the name of the event's action; *target* is the name of the agent that corresponds to the target of the event's action; and *parameters* is a list of optional parameters the event's action might have. For the purpose of illustration, consider a situation in a virtual environment in which agent C invites agent B to a party while agent A hears the invitation. The three agents will perceive the following event $ev_1 < subject = C, action = Invite, target = B, parameters = Party >$. Assuming that agent B appraises this event as very desirable for his goals, then, according to the OCC appraisal theory (Ortony et al., 1988), the agent would potentially elicit an emotion of the type Joy such as $e_1 < type = Joy, valence = 1, intensity = 5, cause = ev_1 >$.

3.1. Empathic Appraisal

Research in neuroscience has shown that parts of the neural network that is involved when one experiences a certain emotional state, such as disgust (Wicker et al., 2003) and pain (Singer et al., 2004; Botvinick et al., 2005), are also activated by the mere observation of that emotional state in another person, regardless of whether the people observed are strangers or loved ones. Based on these notions, the *Empathic Appraisal* phase of our proposed model is triggered either when the agent observes a change[5] in the other agent's *Emotional Cues* in response to an *Event* or by a self-projection appraisal of the event, when the event happens to be the only source of information that the empathizer has. Emotional cues can theoretically be any signal that is perceived by the agent's sensors that indicates the presence of an emotion, such as a facial expression, body posture, or voice tone. After the empathic appraisal is triggered, the *Potential Empathic Emotion* must be determined. As such, the agent must infer what the other agent is feeling. This action is accomplished by combining two different mechanisms: (1) *Emotion Recognition* and (2) *Self-Projection Appraisal*.

The *Emotion Recognition* is used to determine a set of *Candidate Emotions* that are congruent with the emotional cues exhibited by the other agent. Of these candidate emotions, one should be set as the default by the *Emotion Recognition* component, which represents the emotion that is more

[5]Between agents in a virtual environment, the communication of external changes (e.g. changes in the facial expression) is handled by using particular actions that are performed spontaneously (e.g., "Agent A smiles") and are perceived as any other regular event. These actions also include, as an additional parameter, the cause event who triggered them if applicable.

strongly related to the emotional cue itself. Naturally, this process in humans is very complex because humans can consider a vast number of different cues (sometimes contradictory) to infer the emotions of others. To help with this task, humans are also capable of assuming the perspective and the role of others (Feshback, 1978).

Note that, in the proposed model, the *Emotion Recognition* component is intentionally defined in a broad manner. The reason is that, although there are many possible ways to perform such a task, with varying degrees of complexity, the Empathic Appraisal process is not dependent on any specific method. The only requirement imposed by the model is that this component is capable of returning a list of *Candidate Emotions* based on a set of *Emotional Cues* and is capable of defining one of the emotions returned as the default one.

In addition to the *Emotion Recognition* component, the *Self-Projection Appraisal* works as an additional mechanism to infer the emotional state of the empathic target. This process constitutes appraising the event that caused the emotional cues but assuming the other agent's situation. There is, however, an important simplification: the agent uses its own appraisal mechanism and does not consider that other agents could appraise situations differently, because, for example, they have different goals. Consider again the event of agent C inviting B to a party. In this case, the *Self-Projection Appraisal* of agent A toward B is done by simulating the appraisal of an imaginary event where C invited A to a party ($< subject = C, action = Invite, target = A, parameters = Party >$). In this simulated appraisal, A will consider only its own beliefs and goals.

Once the set of candidate emotions ($\mathbb{C} = \{c_1, ..., c_n\}$) and the elicited emotion (ele) are determined, the type of the potential empathic emotion (ee_{type}) is set in the following manner (where c_{def} is the default candidate emotion):

$$ee_{type} = \begin{cases} Type(ele) & \text{if } \mathbb{C} = \{\} \vee \\ & \exists c_i \in \mathbb{C} : Type(el) = Type(c_i) \\ & \\ Type(c_{def}) & \text{otherwise} \end{cases} \tag{1}$$

Using this equation, the type of the potential empathic emotion is equal to the type of the elicited emotion in two situations. The first situation is when there is no observable emotional cues and, thus, no candidate emotions ($\mathbb{C} = \{\}$). This arrangement means that the agent can only imagine what the other is feeling. The second situation occurs when the type of the elicited emotion matches the type of at least one of the candidate emotions. This arrangement means that the self-projection appraisal is in line with what is being recognized from the empathic target's emotional state. When this scenario does not occur, i.e., where none of the existing candidate emotions match the elicited emotion, the type of the potential empathic emotion is equal to the type of the default candidate emotion, as shown in the second branch of the equation.

To provide an example, in the previous party invitation scenario, agent A might not like to attend social parties. As such, he may elicit a Distress emotion, which is not congruent with the smile on agent B. In this case, the potential empathic emotion selected would be Joy, the default candidate emotion for a smile. However, if agent A has never been invited to a party before, it could possibly elicit Pride as the strongest emotion. In this situation, because Pride is one of the candidate emotions inferred from a smile, agent A will assume that agent B is feeling Pride. If the previous scenario is slightly changed so that agent C invites agent B by phone (agent B is not physically present), then agent A can not directly observe any emotional cues. It can only use its own Self-Projection Appraisal to infer the emotion felt by agent B.

After determining the type of the potential empathic emotion, its initial intensity (ee_{ii}) is determined in a similar manner, as shown in (2). In this equation, $I(ele)$ corresponds to the intensity of the elicited emotion, which is determined by the self-projection appraisal process. $I(c_{def})$ is an estimation of the intensity of the default candidate emotion, which is determined by the emotion recognition component. If the type of the elicited emotion (ele) exists in the group of possible candidate emotions then the initial intensity of the empathic emotion is equal to $I(ele)$. If not, the estimated intensity value for the default candidate emotion is used instead as the initial value.

$$ee_{ii} = \begin{cases} I(ele) & \text{if } \mathbb{C} = \{\} \vee \\ & \exists c_i \in \mathbb{C} : Type(ele) = Type(c_i) \\ & \\ I(c_{def}) & \text{otherwise} \end{cases} \tag{2}$$

As depicted in the model's diagram (Figure 1), before the potential empathic emotion is added to the agent's emotional state, its intensity is first modulated by a set of factors. The rationale for this decision relies on the fact that humans do no constantly empathize with every emotion that is perceived in others (de Vignemont, 2006). As stated in (de Vignemont and Singer, 2006) "In real life, we constantly witness people displaying contradictory emotions. If we were to consciously feel what they feel all the time, we would be in permanent emotional turmoil."

To explain why humans only have an empathic emotion in certain situations, de Vignemont and Singer, based on recent neuroscientific evidence, propose a contextual approach for empathy (de Vignemont, 2006; de Vignemont and Singer, 2006). In this approach, empathy is modulated by appraisal processes and "not merely the consequence of the passive observation of emotional cues." (de Vignemont and Singer, 2006) This modulation usually occurs at a fast and implicit level; however, it can also be voluntary, such as in the case of a medical practitioner that is trained to control his own emotions.

In our model, the intensity of the empathic emotion felt by the agent is modulated by four different factors: (1) Similarity; (2) Affective Link; (3) Mood; and (4) Personality. As stated in (de Vignemont and Singer, 2006), the first two concern the relationship between the empathizer and the observed agent, whereas the remaining ones concern the empathizer itself.

3.1.1. Similarity

In 1975, a study conducted by Krebs (Krebs, 1975) showed that subjects that were led to believe that they were similar to someone in personality and values had a stronger empathic emotion than those who were led to believe that they were dissimilar. These findings support the idea that similarity plays an important role in modulating the intensity of the empathic emotion because "we empathize more with people we can identify to" (de Vignemont, 2006).

Based on these notions, similarity (MF_{Sim}) is represented in our model by how much the agent identifies itself with the other agent, in terms of their emotional response to the same event. This amount is determined by the degree to which the emotion elicited by the self-projection appraisal (ele) is similar to the potential empathic emotion (ee):

$$MF_{Sim} = \frac{1}{2} * [S_T(ele, ee) + \left(1 - \frac{|V(ele) * I(ele) - V(ee) * I(ee)|}{I_{Max}}\right)] \quad (3)$$

There are two main components that have the same importance in the similarity equation. The first component looks at the types of the emotions that are being compared; if they are of the same type, then they will be considered similar. The function S_T, which compares the emotion types is defined by:

$$S_T(e_1, e_2) = \begin{cases} 1 & \text{if } Type(e_1) = Type(e_2) \\ -1 & \text{otherwise} \end{cases} \quad (4)$$

The second component of the similarity equation checks the distance between the intensities and valences of both emotions to see how similar they are. $I(e)$ returns the intensity of emotion e, and $V(e)$ returns the valence of emotion e (-1 if negative, and +1 if positive). The intensity distance was normalized to return a value between 0 and 2 and therefore the similarity range is [-1,1]. Note that because the two components of the equation have the same importance, the emotions of different types will never be considered to be more similar than the emotions of the same type. It is important to mention that the notion of similarity employed here is much more narrow than the notion used in (Krebs, 1975) because it focuses exclusively on the emotional response to events. As such, the similarity that exists between two agents can dramatically change from one event to another, without taking into consideration the previous history between them. Still, having this simple notion of similarity is necessary for having agents who do not know each other, thus have no affective link established, still be able to feel more empathy toward one another if they both appraise a situation in a similar way.

3.1.2. Affective Link

Another important modulation factor of the empathic emotion is the affective link (MF_{AfL}) between the empathizer and the person observed (de Vignemont and Singer, 2006). In our model, this concept is represented by a value that symbolizes the social bond that the empathic agent has with the agent observed, namely how much he likes and cares for him. This bond is unidirectional, which means that the value that agent A attributes to agent B can be different than the value that B attributes to A. The value is normalized to vary between -1 (strongly dislikes) to 1 (strongly likes). Like the similarity factor, the affective link enhances (in the case of a positive value) or decreases (in the case of a negative value) the potential of the empathic emotion.

3.1.3. Mood

In the model proposed in (de Vignemont and Singer, 2006), the empathic emotion is not only modulated by the existing relationship between the empathizer and the empathic target but also by characteristics of the empathizer herself. One of these characteristics is the empathizer's mood. As such, in addition to the similarity and the affective link, which characterise the relationship between the two agents involved, our model also uses the mood of the empathizer agent as a modulation factor in the following manner:

$$MF_{Mood} = \begin{cases} Mood, & \text{if } V(ee) > 0 \\ \\ -Mood, & \text{if } V(ee) < 0 \end{cases} \tag{5}$$

Similar to emotions, mood is an affective state that is either positively or negatively valenced. However, an important distinguishing feature between mood and emotions is their respective duration: emotions are typically brief, lasting from seconds to a few minutes, whereas moods can last for several hours (Ekman, 1994; Goldsmith, 1994). There is also a dynamic interaction between the two constructs. Namely, a series of negative emotions leads to a negative mood, while a succession of positive emotions results in a positive mood. Finally, a negative mood has the effect of potentiating negative emotions, and a positive mood potentiates positive emotions (Davidson, 1994).

To capture the aforementioned notions, our proposed model of mood follows the same approach taken in (Dias and Paiva, 2005a), in which mood is represented by a numerical value that ranges from -1 to 1, with 0 signifying a neutral state. When a negative/positive emotion is added to the emotional state, the mood value decreases/increases. The value then slowly returns to 0 where it stabilizes. With regards to modulating the empathic emotion, a negative mood increases the intensity of an empathic emotion with a negative valence ($V(ee) < 0$) and decreases the intensity on an empathic emotion with a positive valence ($V(ee) < 0$). On the other hand, a positive mood works in the opposite manner. The valence of an emotion is directly associated with its type. For example, in OCC theory, Joy, Pride and Admiration are emotion types that have a positive valence, whereas Distress, Shame and Reproach are emotion types with a negative valence.

3.1.4. Personality

Finally, the last modulation factor included in our model is the personality of the empathizer agent, which is also considered in (de Vignemont and Singer, 2006) to be one of the empathizer characteristics that modulate the empathic response. For instance, people with a personality that is prone to feeling negative emotions are also prone to feeling negative empathic emotions (Eisenberg and Morris, 2001).

Based on these findings, in our proposed model, personality is represented as the agent's resistance to feeling certain emotions, which is defined with a threshold value (*EmThreshold*) that ranges from 0 to 1 for every type of emotion (a similar model of personality is used in (Dias and Paiva, 2005a)). Therefore, empathic emotions to which the agent has a weaker/stronger resistance will be more/less likely to be added to the emotional state. To support the aforementioned facts, a personality modulation factor was defined as follows:

$$MF_{Pers} = -EmThreshold(ee_{type}) \quad (6)$$

It is important to mention that this approach to modeling the effect of personality in emotional appraisals was also explored in (Doce et al., 2010), where the Big Five model of personality (John and Srivastava, 1999) was mapped to different emotional thresholds. In the study conducted by the authors, the results indicate that the chosen mapping was capable of altering the user's perception of the agent's personality traits in a significant manner.

3.1.5. Empathic Emotion

Each one of the modulating factors contribute in the following manner to the intensity of the empathic emotion (ee_{fi}).

$$ee_{fi} = \min\{ee_{ii} * [MF_{AfL} + MF_{Sim} + MF_{Mood} + MF_{Pers}], MaxI\} \quad (7)$$

Note that in the above equation, $MaxI$ corresponds to the maximum intensity an empathic emotion can have. Also note that the factors in the above equation are independent from each other. For example, even if the similarity value is zero, an empathic emotion can still occur if there is a positive affective relationship and the opposite is also true. For instance , if agent B is a very good friend of agent A (a high affective link) and agent B is overjoyed for something really strange according to agent A (a highly negative similarity), then A will not empathize with B's joy.

The resulting empathic emotion is finally added to the emotional state, but only if the determined intensity (ee_{fi}) is greater than zero. After being added, its intensity will start to decay and fade out with time. As such, it is possible that when adding a new empathic emotion, other emotions are still active in the emotional state. These other emotions could have been generated by previous empathic experiences or by other appraisal processes. As a result, the agent is capable of "feeling" more than one emotion at any given time. This allows the representation of "mixed feelings" that can occur, for instance, when an event is undesirable for a friend but desirable for us simultaneously.

To exemplify the interplay of all of the proposed Modulation factors, let us consider again the same scenario with agent C inviting agent B to a party which raises a smile in B's face. What can happen to Agent A in terms of empathy toward B? When A projects himself into B's situation, if a Joy emotion is elicited, then they have a high similarity with one another. This scenario will likely cause A to feel an empathic Joy emotion. However, imagine that A strongly dislikes B; in other words, A has a negative affective link toward B. In this case, it is more unlikely that A will empathize. Furthermore, if the affective link of A toward B is neutral but A is in a negative mood or has a personality with a strong resistance to feeling Joy, then it is harder for A to empathize with B's joy emotion.

Finally, to further clarify how the *Empathic Appraisal* is accomplished, Algorithm 1 describes in pseudocode the entire process. After a new event is perceived or new emotional cues are observed a list of empathic targets is created by the function $DetermineEmpathicTargets(evt, ec)$. This list will contain both the event's subject and target if they are different from the empathizer agent, which is referred to as $SELF$. In addition, the list will also contain the agents that had a change in their emotional cues. For each empathic target, a set of candidate emotions ce is obtained by the call to the $EmotionalRecognition(empTarget.ec)$ function. The following "if-then" block corresponds to the *Self-Projection Appraisal* which is performed only in case in which there is an event ($evt \neq null$). It starts by creating a copy ($sevt$) of the event and then substituting the empathic target with $SELF$ (the empathic target can be either the event's subject or the event's target).

Algorithm 1 Empathic Appraisal

$evt \leftarrow PerceiveEvent(); ec \leftarrow PerceiveEmotionalCues()$
$empTargets \leftarrow DetermineEmpathicTargets(evt, ec)$
for all $empTarget \in empTargets$ **do**
  $ce \leftarrow EmotionalRecognition(empTarget.ec)$
  if $evt \neq null$ **then**
    $sevt \leftarrow evt.copy()$ {"sevt means simulated event"}
    if $evt.subject = empTarget$ **then**
      $sevt.subject \leftarrow SELF$
    else
      $sevt.target \leftarrow SELF$
    end if
    $ele \leftarrow Appraisal(sevt)$
    $MF_{Sim} \leftarrow DetermineSimilarity(ele, ce)$
  else
    $ele \leftarrow null; MF_{Sim} \leftarrow 0$
  end if
  $ee.type \leftarrow DetermineEEType(ele, ce)$
  $ee.intensity \leftarrow DetermineEEIntensity(ele, ce)$
  $MF_{AfL} \leftarrow getAfLink(empTarget);$
  $MF_{Pers} \leftarrow getThreshold(ee.type)$
  $MF_{Mood} \leftarrow determineMoodFactor()$
  $ee.intensity \leftarrow ee.intensity * [MF_{AfL} + MF_{Sim} + MF_{Mood} + MF_{Pers}]$
  if $ee.intensity > 0$ **then**
    $AddEmotionToEmotionalState(ee)$
  end if
end for

Afterwards, it uses the existing agent's appraisal method on the simulated event to determine the elicited emotion. Subsequently, the similarity factor (MF_{Sim}) is determined. If no event was perceived, then this factor is set to zero. After the *Self-Projection Appraisal* is performed, the potential empathic emotion is created. Its type is determined by equation 1 and its initial intensity by equation 2. Subsequently, the remaining modulation factors ($MF_{AfL}, MF_{Pers}, MF_{Mood}$) are calculated using their respective equations. These factors are then used to modulate the intensity of the empathic emotion. Finally, if this intensity is higher than zero, the emphatic emotion is added to the emotional state.

3.2. Empathic Response

Hoffman defines empathy as "an affective response more appropriate to another's situation than one's own"(Hoffman, 2001). In light of this definition, our model considers that an empathic response necessarily involves an emotion generated by the empathic appraisal. This view is also supported in (de Vignemont and Singer, 2006), where it is argued that empathy exists only if the empathizer is in an affective state. An empathic emotion felt by the agent can then potentially trigger a specific empathic action. As an example, consider the act of congratulating agent B, which is triggered by agent A feeling Joy for agent B after B obtains a party invitation. In our model, empathic actions such as the one just mentioned are defined by the following properties (similar to the properties used in (Paiva et al., 2004) for action tendencies):

- **Empathic Action** - the name of the empathic action that is triggered;
- **Eliciting Empathic Emotion** - the specific eliciting empathic emotion that triggers the action;
- **Cause Event** - the event that caused the empathic emotion.

Algorithm 2 describes the process of selecting an empathic action. Overall, it goes through the list of every empathic action defined in the agent's profile and attempts to see whether the eliciting empathic emotion of any action matches one of the empathic emotions that the agent currently has in its emotional state. If a match is found and if the emotion has sufficient intensity there is an additional check concerning the cause event. Then, in the end, the selected action will be the one with the highest emotional intensity (or none if no match was found).

Algorithm 2 Empathic Response

```
empEmotions ← GetCurrentEmpathicEmotions()
empActions ← GetEmpathicActions()
selAction ← null
for all a ∈ empActions do
  for all e ∈ empEmotions do
    if e.type = a.elEm.type and e.intensity > a.elEm.intensity then
      if a.causeEvent ≠ null and a.causeEvent Matches e.causeEvent
      then
        if selAction = null or e.intesity > selAction.intensity then
          selAction ← a
          selAction.intensity ← e.intensity
        end if
      end if
    end if
  end for
end for
return selAction
```

4. Integration into an Affective Agent Architecture

The proposed model was integrated into an affective agent architecture, named FAtiMA, (Dias and Paiva, 2005a) that is capable of generating emotions to strongly influence the behavior of synthetic characters. In this architecture, the emotions are synthesized by an appraisal of events that stems from the OCC cognitive theory of emotions (Ortony et al., 1988). As such, it was possible to simply reuse the existing appraisal process to implement the self-projection appraisal of our empathy model. Note that the same could be done with any other existing emotional architecture that models an appraisal process, such as EMA (Marsella and Gratch, 2009), even if such process is based on a appraisal theory that is different from the OCC. Another characteristic of FAtiMA that is important for the purposes of our model is that it features a dynamic mood model and it allows the manual specification of emotional thresholds as a way to define a character's personality profile. This means that two of the modulation factors proposed in our model, mood and personality, were already present, minimizing the implementation effort. Figure 2 shows the final agent architecture.

In this architecture, events are perceived from the environment using

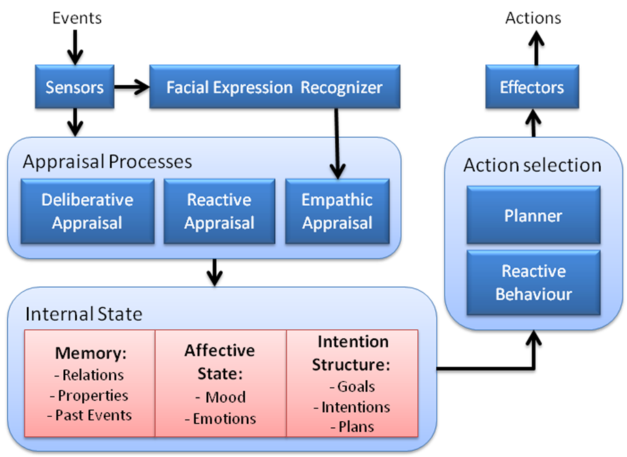

Figure 2: Empathic Agent Architecture

the sensory apparatus of the agent. Given that our case study takes place in a virtual environment with virtual agents only, the emotional cues used were simple logical properties such as "$FacialExpression(John) = Smile$". These properties are directly perceived by every agent that is present in the scene. Thus, to apply the model in a situation in which agents interacted with humans, a more complex mechanism of detecting emotional cues would be required, for instance, real time facial tracking of the user's face. However, this added complexity would not change the structure or nature of the empathic process that is proposed in this model.

Given that the facial expression of other virtual agents are directly observable properties, i.e., there is no detection error, the process of emotion recognition becomes straightforward. Specifically, a predefined mapping between facial expressions and specific emotion types is applied. When an agent changes his expression, every other agent in the environment is notified. Again, this simple approach was used because the focus of this work is not on emotional recognition. A more complex approach would be required to have agents empathizing with users.

Regarding the appraisal processes, the already existing deliberative and reactive appraisals were retained. The first appraisal elicits prospect-based emotions that are associated with the relation between the events and the agent's goals. These emotions are used to influence the agent's deliberation and planning. On the other hand, the reactive appraisal elicits all of the other types of emotions by using predefined reaction rules, which are specific to each character. These emotions are used by the reactive behavior component to generate quick emotional reactions.

Our added empathic appraisal works concurrently with the other appraisal processes. When an event is perceived, the agent projects himself as if he was the agent who triggered the event and uses a simulated version of his own reactive appraisal for that projection. From this simulation, an emotion is elicited, which is then compared with the candidate emotions that are associated with the agent's facial expression, to create a new potential empathic emotion.

The similarity factor is then calculated by comparing the elicited emotion with the potential empathic emotion. On the other hand, the affective link value corresponds to the "like" relationship value that is specified for each agent in relation to every other agent.

Before the emotion is added to the emotional state, its base intensity is affected by the determined similarity and affective link and also by the character's current mood and personality specification. This entire process is then repeated for the other agent to whom the event was directed. The empathic emotions that are added to the affective state of the character are then used by the reactive behavior component to generate empathic actions, based on the action rules that were predefined for them.

5. Creating Empathic Agents

To illustrate how the model can be used to generate empathic emotions in an emergent manner, we developed a simple multi-agent scenario that aims to illustrate a bullying situation. The scenario developed is similar to the scenarios found in the FearNot! application (Paiva et al., 2005), from which we reused the same framework and graphical assets.

The scenario involves five different characters (see Figure 3), in which each of the character is an autonomous agent driven by the architecture described in the previous section. As such, instead of following a script, the behavior in the scenario emerges from the character's individual goals and emotional parameterization (see Table 1), which are based on the intended role that each character has in the scenario. John plays the victim role and Luke plays a bully. Together, they are the main characters. Ollie, Paul and

Figure 3: The characters in the schoolyard

Frances have secondary roles: Ollie is a friend of John, Paul plays the role of another bully that dislikes John and Frances is an emotionally neutral character that mostly observes the interaction.

Table 1: Agent Emotional Parameterization.

Agent	Reactive Appraisal Rules			Affective Link [-1..1]	Personality (Emotional Threshold) [0..1]
	be complimented	insult another	be insulted		
	Desirability [-10,10]	Desirability [-10,10]	Desirability [-10,10]		
Luke	3	6	-6	-0.5 for John	0
John	5	-5	-6	0 for all	0
Paul	3	4	-6	-0.5 for John 0.5 for Luke	0
Ollie	5	-5	-6	0.5 for John	0
Frances	3	-3	-6	0 for all	1

Using the same goals and emotional parameterization, two versions of the scenario were created: (1) characters with self-oriented emotions only, and (2) characters with both empathic and self-oriented emotions. Table 2 describes the resulting behavior in both cases.

Table 2: Interaction Sequence*

Frances:	Hi guys! Were you playing football again?
Ollie:	Yes.
Paul:	It was fun.
Luke:	I think John was the best player.
John:	(smiles) Thank you Luke!
Ollie:	**(smiles)**
Luke:	I was joking you idiot. You were really awful. (smiles)
John:	(sad face) Why are you always making fun of me?
Paul:	**(smiles) Nice joke Luke!**
Ollie:	**(sad face) Nevermind him John! He is just a jerk.**

* The text in bold refers to the behaviors that appear only in the version with the empathy model. The non-bold text refers to the behaviors that occur in both versions.

5.1. Self-Oriented Emotions

In both versions, the characters are capable of simulating self-oriented emotions, i.e., emotions that relate to their own situation in opposition to empathic emotions that concern the situation of another. Characters can simulate such emotions by appraising the events that occur in their environment, using their reactive appraisal rules, which are shown in Table 1. Such rules are part of the FAtiMA appraisal model, and their formalization is described in (Dias and Paiva, 2005a). These rules allow us to specify how certain appraisal variables from OCC theory, such as Desirability, are determined in response to certain events. In this case, the events are: (1) receive a compliment, (2) insult another, and (3) receive an insult.

Given that the values specified in these rules are not always identical for all of the characters, it is likely that the same event triggers different reactions. In this scenario, John finds the act of being complimented to be desirable. As such, when Luke compliments him, he generates an emotion of joy, which makes him smile. Subsequently, Luke insults John, something he feels joy for doing because, in his perspective, it is a desirable thing to do.

However, John feels distressed for being insulted, because he, similar to all of the characters, considers it to be an undesirable event.

5.2. Empathic Emotions

Aside from the self-oriented emotions, the characters also have empathic emotions in the second version of the scenario. These emotions emerge as a result of the use of our model in the following manner. When John smiles after being complimented by Luke, every other character including Luke registers this smile as an emotional cue and puts themselves in John's place. They then simulate how they would feel if they were complimented. In this case, because every character finds the act of being complimented as desirable (see Table 1), all of them would feel a joy emotion in the situation of John, an emotion that is congruent with his smile. As such, the degree of similarity between John is high for every character. Therefore, based on similarity alone, every character would feel joy for John in this situation. However, this result is not what occurs given the other modulation factors that are considered in our model. As shown in the interaction described in Table 2, it is only Ollie that has an empathic emotion. Luke and Paul do not empathize in this situation because they have a negative affective link toward John. In the case of Frances, she does not have an empathic emotion as well, but because of her emotional threshold, which is very high.

The second situation in which an empathic emotion arises is when Luke insults John. The emotional cues triggered by this event are a smile on Luke's face and a sad expression on John's. The empathic process is then triggered by every other character toward both Luke and John. In the case of Luke, the others will appraise how they would feel if they insulted John. In this case, Paul is the only one that finds it desirable, which means that the similarity factor will be high. Oppositely, the similarity factor is low in the case of Frances and Ollie because they find it undesirable to insult others. Consequently, Paul is the only character that has an emphatic emotion toward Luke, which makes him smile (see Table 2). In the case of John, the other characters will appraise how they would feel if they were insulted. All of them would consider it to be undesirable. Still, similar to what happened with the compliment, only Ollie empathizes with John. The empathic emotion of Ollie then triggers an empathic action toward John, which corresponds to a speech act of consolation (see Table 2).

6. Evaluation

The purpose of the previous scenario was to illustrate how the proposed model of empathy is capable of endowing agents with the ability to empathize with others. As shown in the scenario, the model generates different empathic emotions in response to the same events. This result occurs from the fact that agents have different emotional thresholds, different affective links to one another and they use their own appraisal mechanism to judge how similar they are. Consequently, a striking difference in the empathic emotion is shown when John is insulted by Luke. Specifically, Paul feels happy for Luke and Ollie feels sad for John. The goal of this empirical evaluation is to investigate what are the psychological effects of the empathic behavior produced by the proposed model on the user's opinion about the agents. A similar goal was pursued in the empirical evaluation of the empathy model proposed in (Boukricha et al., 2013). This evaluation will be discussed in the Related Work section.

6.1. Hypotheses

One of the first studies to empirically investigate the psychological effects of agents with empathic emotions upon people was reported in (Brave et al., 2005). Their findings suggest that, compared to agents with self-oriented emotions only, agents with empathy are perceived as being more caring, likable, trustworthy and submissive but no significant difference exists concerning intelligence or dominance. If our proposed model is capable of endowing agents with the ability to empathize with others, then it appears to be reasonable to expect similar results when the model is applied compared to when it is not. However, it is important to consider that the study conducted by Brave and colleagues evaluated only the effects of empathy when it is directed toward the user and his or her situation. Instead, in our scenario, agents empathize with other agents that are present in the same virtual environment, socially interacting with one another. In this case, the interaction revolved around a bully insulting a victim. Because most people find bullying to be morally wrong, we expect that the results obtained in Brave's study will apply only to the empathy expressed by Ollie (the victim's friend). In the case of Paul, who empathizes with Luke (the bully character), we expect to obtain the opposite effect. Finally, as Frances does not empathize with anyone in this situation, there should be no effect on how users perceive her. In summary, our goal is to test the following hypotheses:

- H1: With the empathy model, participants perceive Ollie as being more caring, likable, trustworthy and submissive.
- H2: With the empathy model, participants perceive Paul as being less caring, likable, trustworthy and submissive.

6.2. Design and Procedure

Two videos were recorded of the system running, one in which the characters had only self-oriented emotions and another in which the characters also had empathic emotions generated by the proposed model. The videos were then used in an online questionnaire in a within-subjects experiment, where the participants viewed both of the videos in a random order. As a result, roughly half of the participants viewed the video with the empathy model first, and the other half viewed the video without the empathy model first. After seeing each video, the participants had to answer a set of questions about it.

The subjects were asked about which characters they liked and disliked and also whether they felt sorry for John or happy for Luke (using 7-point Likert scales). The participants then had to rate their opinion about whether the three other agents, Frances, Ollie and Paul, were: Caring, Likeable, Trustworthy, Intelligent, Dominant, and Submissive. All of these items were measured directly using a 7-point Likert scale for each. Finally, we asked the subjects their gender and age.

6.3. Results

We had a total of 77 participants, with an average age of 27 years old, of which 19 were female and 58 were male. The majority of participants were Portuguese. Given that multiple comparisons were performed ($n = 20$), the criteria for statistical significance used was $p < 0.0025$ instead of $p < 0.05$. This new value was determined by applying the Bonferroni correction.

6.3.1. Results - Character Preference

The results obtained for the users' favorite character in both conditions are shown in Figure 4. As shown, Ollie goes from being one of the least favorite characters (9%) to becoming the most favorite character in the empathy condition, where he is chosen by 60% of the users. The McNemar chi-square test for independence was applied to determine whether this difference obtained concerning the preference about Ollie was significant or not.

To apply the test, participant responses were grouped in two categories: (1) Preferred Ollie, (2) Preferred another character. The result obtained was strongly significant ($p < 0.001$). A possible explanation for this result is that the empathy model helped to portray Ollie's friendship and concern with John. This effect, in turn, made viewers relate more with Ollie.

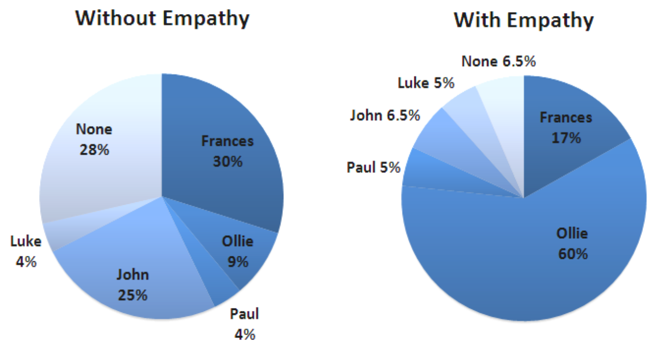

Figure 4: Question: "Which character did you like the most?"

Regarding the question "Which character did you dislike the most?", in both cases, Luke was the character who was the most disliked, selected by 53% of the users with the empathy model and by 74% without the model. In the empathy model condition, Paul was the second most disliked character, chosen by 23% of the users. Interestingly, Frances was very popular in both conditions, being chosen as the most favourite character in the video without the empathy model and scoring second place in the other video. However, it is difficult to explain this result because the only relevant differences between Frances and the other characters are the fact that she is a girl and the fact she was emotionally neutral in both conditions.

6.3.2. Results - Empathy Felt by the User

Figure 5 shows the results regarding the user's empathy towards John and Luke. Users empathized with John (they felt sorry for him) but not with Luke (they did not feel happy for Luke). The Wilcoxon statistical test was applied to determine the impact of the empathy model and the results were not significant in both cases ($p = 0.54, p = 0.715$). A nonparametric test was used because the data for the two variables did not follow a normal distribution.

Figure 5: Results for the empathy felt by the user. The error bars represent the standard deviation.

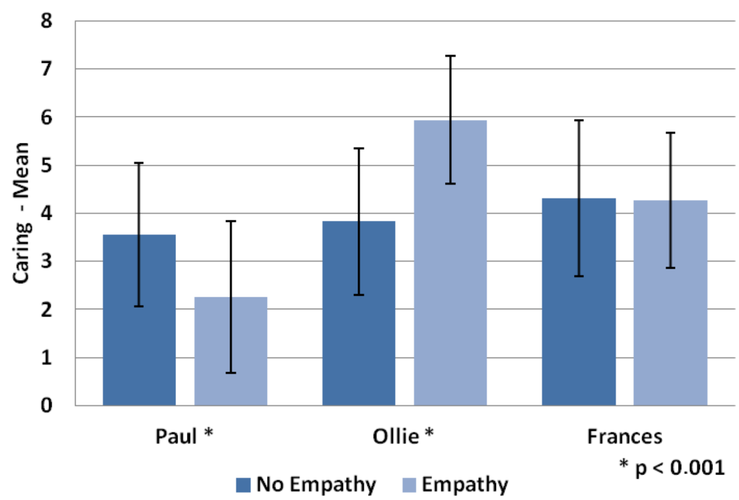

Figure 6: Caring results. Error bars represent standard deviation.

6.3.3. Results - Caring

Figure 6 shows the results obtained in terms of how much the users agreed in describing Paul, Ollie and Frances as caring. Once more, the Wilcoxon test was used to compare the two conditions (the obtained distributions were significantly different from the normal distribution). The same approach also applies for the other characteristics analyzed in this paper.

As shown in the graph, there was a significant effect on both Paul and Ollie, yet in opposite directions. With the empathy model, Paul was significantly perceived as being less caring ($z = -5.31, p < 0.001$) but Ollie was perceived as being more caring ($z = -6.33, p < 0.001$). The effect size was medium for Paul ($r = 0.42$) and high for Ollie ($r = 0.51$)

6.3.4. Results - Likeability

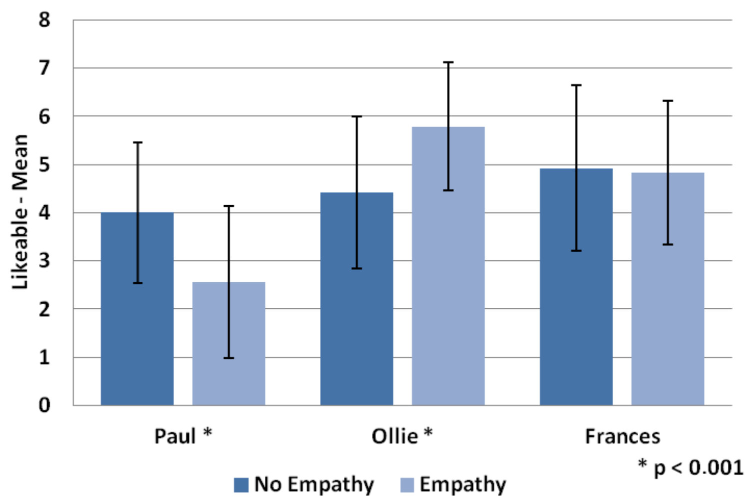

Figure 7: Likeability results. The error bars represent the standard deviation.

The results obtained for the characters' likeability are shown in Figure 7. There were significant effects for Paul and Ollie but not for Frances. Paul decreased his perceived likeability ($z = -4.59, p < 0.001$) in the empathy condition, whereas Ollie significantly increased it ($z = -4.48, p < 0.001$). The effect sizes were both smaller than the effects obtained for caring, but they are still medium size effects, ($r = 0.37$) in the case of Paul and ($r = 0.36$) in the case of Ollie.

6.3.5. Results - Trustworthiness

The results for the perceived trustworthiness are depicted in Figure 8. As the graph shows, the results are very similar to the ones obtained for caring and likeability. Paul, under the empathy condition, is described as significantly less trustworthy ($z = -4.76, p < 0.001, r = 0.38$) while Ollie increased his trustworthiness ($z = -5.68, p < 0.001, r = 0.46$).

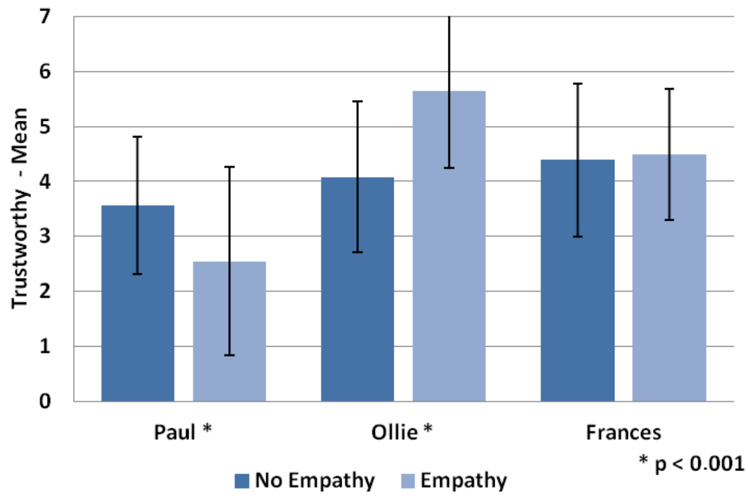

Figure 8: Trustworthiness results. The error bars represent the standard deviation.

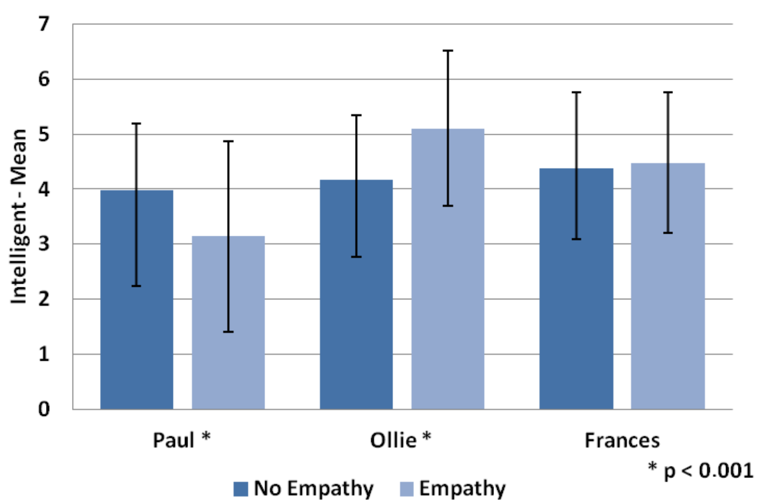

Figure 9: Intelligence results. The error bars represent the standard deviation.

6.3.6. Results - Intelligence

Figure 9 shows the result for the users' agreement in describing Paul, Ollie and Frances as intelligent. On the one hand, Paul was significantly perceived as being less intelligent under the empathy condition ($z = -4.43, p < 0.001, r = 0.36$). On the other hand, Ollie was described as being more intelligent with the empathy model ($z = -4.36, p < 0.001, r = 0.35$). There was no significant effect for Frances.

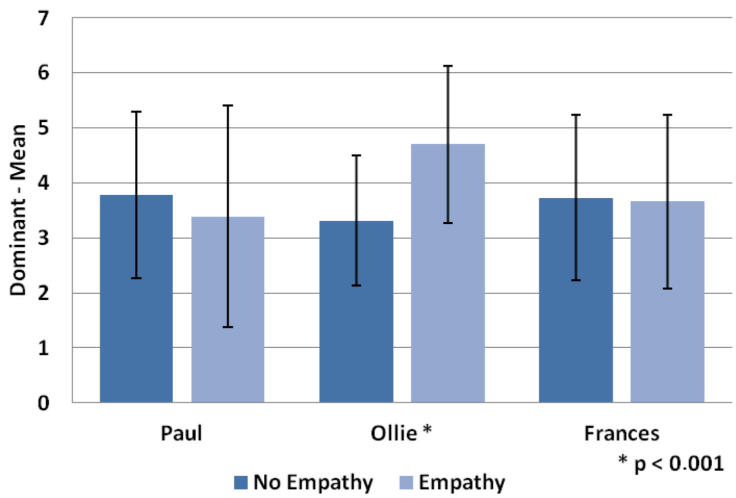

Figure 10: Dominance results. The error bars represent the standard deviation.

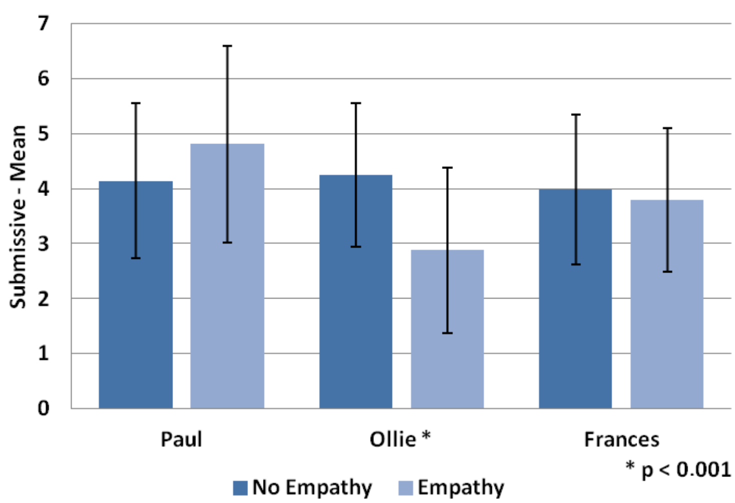

Figure 11: Submissiveness results. The error bars represent the standard deviation.

6.3.7. Results - Dominance

The results for dominance are depicted in Figure 10. As shown, Ollie is the only character that has a significant difference. Specifically, his perceived dominance was higher in the empathy condition ($z = -5.414, p < 0.001, r = 0.44$).

6.3.8. Results - Submissiveness

Figure 11 shows the results that were obtained for the final characteristic analyzed, which was submissiveness. Again, Ollie was the only character that had a significant difference. He was perceived as being less submissive ($z = -5.14, p < 0.001, r = 0.41$) in the video with the empathy model.

6.4. Discussion

The results that were obtained for the characters' attributes are summarized as follows. While Frances was always perceived in a similar manner in both conditions, the participants changed drastically their opinion about Ollie and Paul. Ollie had significant effects on all of the six characteristics. Namely, with the empathy model, users thought he was more caring, likeable, trustworthy, intelligent, dominant, and less submissive. The strongest effect was associated to caring ($r = 0.51$). Paul was significantly perceived as being less caring, likeable, trustworthy and intelligent with the empathy model.

Overall, these results further indicate that empathic agents are perceived differently from those that are not empathic. However, the effect of having empathy is not always the same. In fact, the effect can be the opposite, as in the case of Ollie and Paul. Ollie, similar to the results obtained in the study conducted in (Brave et al., 2005), was perceived as being more caring, likeable and trustworthy. This perception occurred even though he was empathizing with another character and not with the user as in the aforementioned study. Note that the vast majority of users felt empathy toward John, the victim. In this sense, both Ollie and the users had a similar empathic response, which in turn could have originated a similar effect as if Ollie was empathizing with the user as well. This consideration can also explain why the users felt that Paul, who empathized with the bully instead, was less caring, likeable and trustworthy.

An important difference between our study and the one conducted in (Brave et al., 2005), is that there was no significant effect for the perceived intelligence of the empathic agent in the latter, whereas in our study, Ollie was perceived as more intelligent and Paul as less intelligent, in the empathic condition. Note that the activity performed by our agents in both conditions is purely based on social dialogue, in contrast to the game activity (playing blackjack) of the other study. As such, it could be the case that users focused on the socio-emotional intelligence of the agents in our scenario and on their capacity for playing blackjack in the other scenario, which did not vary. Finally, another important difference between the two studies was that Ollie was significantly seen as being more dominant and less submissive, whereas in (Brave et al., 2005), the empathic agent was seen as more submissive, and there was no significant effect for dominance. This result is likely because Ollie stood up to a bully after he felt sorry for John. This result further indicates that the nature and context of the situation in which empathy occurs, as well as to whom the empathy is directed, are all important factors in the judgement of the observed empathic behavior.

6.5. Limitations

One of the difficulties in validating the proposed model arises from the fact that it attempts to simulate internal processes inspired by psychological models of human empathic behavior. Furthermore, as argued by Aylett and Paiva concerning this type of evaluation (Aylett and Paiva, 2012), "there is no straightforward way to test a computational model like this one in isolation from a concrete application and scenario". As such, to assess the psychological impact that empathic agents have on users, we must contextualize the interaction and build a concrete scenario and application. Further, for a human observer to judge such "modeled" processes, they must be externalized in some fashion. Thus, to create a test scenario, we had to make decisions about how such externalization was to be constructed. This involved, for example, deciding which exact sentence is said by Ollie when he is having an empathic emotion caused by the proposed model which results in an empathic action. Thus, we had to make the "actions" very concrete, although they result from the internal processing of the model, and with different modulating factors, other actions would emerge as a result of the model. However, one should also acknowledge, that in the experiment conducted, it is possible that such decisions on what actually is said influenced the results. As such, it is hard to know the extent that the results obtained are generalizable to other scenarios or to other characters without further testing. Nevertheless, these types of choices are always necessary as soon as we place the model in a concrete environment, and especially in a virtual agent scenario. Having carefully considered these issues, we attempted to minimize confounding factors. Nevertheless, we agree that a single experiment, such as the experimented performed, limits any strong conclusions that can be drawn about the validity of the model, and further studies must be conducted. For instance, it is important to determine the individual contribution of the different modulation factors that were implemented in the model. In this experiment the model was tested in its entirety, so it is not possible to know if all of the factors are equally important and if our assumptions in their implementation are validated.

Another limitation of this study lies in the type of measures that were adopted in the evaluation. In most of the work studied, the evaluation of empathic agents is usually performed by assessing the impact of the agent in its "relation" established with the user. For example, in the work by Ochs (Ochs et al., 2012), an agent is evaluated in terms of the user's perception of many of its characteristics, in particular its facial expressions and social aspects, such as pleasant, expressive, cold, jovial, boring, strict, and cheerful. The work by Brave et al. (Brave et al., 2005), which served as inspiration for this concrete evaluation, also relies on such type of assessment. In spite of this consideration, one must acknowledge that evaluating virtual agents through subjective questionnaires is influenced by many small details that emerge from the behavior of the agent. For example, recently Krämer et al. (Krämer et al., 2013) showed that just the simple fact that a virtual agent smiles at the user leads to a different perception of that agent. In fact, the work presented here, we rely on these subtle effects in the user by making the behavior of the agent (the agent's empathic responses dictated by the model that determines how much it smiles and at whom) dependent on the empathic processes that were modeled. Yet, other psycho-physiological measures, such as the heart rate (HR) and electrodermal activity (EDA), because they do not imply subjective questionnaires, would most likely provide more reliable information on the emotional reactions of the users to the agents. However, because these measures are good indicators of arousal but not of valence, and they do not tell us anything about the perception of the relation toward the agent, these measures would not provide us with sufficient information for what we aimed to accomplish. A combination of physiological and self-report measures would be interesting, and recently Rosenthal-von der Pütten et. al. (Rosenthal-von der Pütten et al., 2013) have such a combination, which they used to study the empathic responses of users to robots, achieving very good results. However, because we wanted to see the impact that the "empathy model" had on the generation of different individual emotional behaviors, our goal was not as much to evaluate the empathy felt by the user.

Finally, one can further question why, in a study that is focused on empathy, there is no empathy questionnaire used in its evaluation. Most of the empathy questionnaires, such as the IRI (Davis, 1983) or the EQ (Baron-Cohen and Wheelwright, 2004), were designed to identify individual differences in humans on what concerns empathic reactions to situations. However, our focus is not to assess the empathy that is felt by the user; instead, we attempt to assess the impact that the empathy model of the agent has on other factors involving the user, to create individual believable "empathic" agents. As such, an empathy questionnaire applied to the users would add another dimension to the problem. That usage would certainly expand our understanding of the impact that our agents have on different users, but it would not help in the concrete evaluation of the model itself, in particular in its capability of generating different empathic behaviors that users perceive as such. We, however, hope that, in the future, this new dimension can be added to further studies and scenarios.

7. Related Work

In this section, we will relate the research presented here to several other existing models in which empathy is addressed and partially captured. Perhaps one of the first developments in this area was the seminal work by Elliot (Elliott, 1992), who implemented an Affective Reasoner (AR), one of the first implementations of the OCC model of emotions (Ortony et al., 1988). In this work, agents were able to reason about events that involve other agents through a construed mental model of their goals. By using pre-defined types of social relationships, such as "Friendship" and the "Empathetic Unit", the Affective Reasoner was capable of producing some empathic outcomes even though there was no explicit model of the empathic process. In contrast to our model, which focuses on the empathic processes, in Elliot's work, the empathic behavior is a direct consequence of the relationships that are pre-defined between the characters.

Also based on appraisal theory, the work by Ochs et al. (Ochs et al., 2009) considers modulating factors such as the social relations of agents (in this case, Non Player Characters- NPCs in a game) as well as their personality, in the generation of their emotional behaviors. More recently, Ochs et al. (Ochs et al., 2012) focused their work on the dialogue between an agent and the user, and they built an empathic dialog agent that uses a formal representation of emotions to understand the user's emotional state during a dialog. The formal model, which uses a logic framework, characterizes the mechanisms behind empathic dialogues between a virtual agent and a user. In that dialogue, after inferring the user's affective state, the agent elicits an empathic emotion of the same type. The intensity of the empathic emotion is determined by the product of the intensity of the inferred user's emotion with the *degree of empathy* that the agent has for the user. The latter represents a general modulation factor for the empathic emotion. The results of a study conducted with the empathic agent showed that the user's perception of the agent changed significantly by the presence of the empathic behaviors that were generated. In contrast to our proposed model of empathy, in this work, the user's emotional state is inferred only by perspective taking. Instead, in our model the emotion that the user is feeling is inferred by determining the emotion the agent would feel if it was in the user's place but also by looking at the emotional cues that the user is expressing.

The recent work of H. Boukricha (Boukricha and Wachsmuth, 2011), which is very much in line with our work, features an empathy model that was implemented and tested in a virtual human, EMMA, which includes features that are essential for the expressivity of empathic interactions, in particular, emotional facial expressions that are based on a dimensional model of emotions (the PAD-Pleasure Arousal Dominance model). In an innovative manner, by using the PAD dimensions and facial expressions, the model also distinguishes not only the different mechanisms that are involved in empathy processes (in particular facial mimicry and situation role-taking), but also the way that these processes are modulated by internal and external factors of the agent. Further, by using a similar model and the virtual agent MAX, a spatial interaction collaborative scenario was created (Boukricha et al., 2011). There, MAX collaborates with the user, and the more that MAX likes that partner, the stronger the empathic responses become toward the partner. Thus, by endowing some degree of empathy into the virtual human, distinct helping actions between the agent and its human partner emerge. The perception of this degree of empathy was further tested in a subsequent study (Boukricha et al., 2013). In this experiment, participants observed recorded video interactions between EMMA, MAX, and a real person. In these videos, the person would always start by greeting and giving a compliment to EMMA, which triggered a positive emotional response in EMMA. Afterwards, the person in the video would greet and insult MAX, causing a negative emotional response in MAX and triggering an empathic response from EMMA. Three experimental conditions were created regarding the degree of empathy shown by EMMA: (1) neutral, (2) medium and (3) maximum. These were obtained just by manipulating the liking factor of their empathy model, which is similar to the affective link factor of our proposed model. The results obtained showed that participants were able to perceive the different degrees of empathy showed by EMMA and also, they significantly rated EMMA as more likeable in the condition with the maximum degree of empathy. This result in is line with the result obtained in our experiment concerning the perception of the character Ollie who empathized with the victim in the empathic condition.

Another interesting work in empathic interaction is described in (Prendinger and Ishizuka, 2005), in which an empirical approach is explored, which relies on physiological data of the user as the basis for empathic feedback. By implementing a life-like character companion, whose goal was to act as an empathic tutor in a training scenario for job interviews, they explored empathy in the responses to the user. The user's physiological data are gathered and then translated into emotions, which allows the companion to provide affective feedback by using specific empathic responses of comfort, encouragement or praise.

Although there are studies that show that it is more difficult to distinguish emotions along the valence axes than along the arousal axes (Wagner et al., 2005), in the work by Prendinger the empathic tutor is equipped with the ability to capture valence as well as arousal. The emotions that were considered were the following: "relaxed" (happiness), defined by the absence of autonomic signals, i.e., no arousal (relative to the baseline); "joyful" defined by an increased arousal and positive valence; and "frustrated", defined by increased arousal but negative valence. Although this approach appears to be extremely promising, it is based on physiological signals and suffers from some of the known limitations for inferring emotional states from physiological data, namely, the "Baseline Problem", which refers to the problem of finding a condition against which physiological change can be compared with (its baseline); the "Timing of Data Assessment Problem" which refers to the temporal dimension of emotion elicitation; and the question of how the intensity of an emotion is indeed reflected in the physiological data that is obtained.

However, despite these limitations, comparisons made between the empathic and the non-empathic version, showed that users had lower levels of arousal and stress when interacting with the empathic tutor. Indeed, to create agents that respond to users in an empathic manner, it is important to capture (or infer) the emotional state of the user. However, given the vast number of emotional recognition approaches and modalities, each one with their own advantages and disadvantages, we decided to view the recognition of emotions in a generic way. As such, our model does not compromise with any specific emotion recognition modality and method. Instead, it intentionally leaves the process of emotional recognition defined in a very general manner, to allow it to be applied either to virtual agents or to humans interacting with them, which is not possible using physiological data as in the previous work. By being unconstrained, the emotional recognition can be implemented depending on the specific domain in which the model is to be applied.

Along the same lines of investigation as Predinger and Ishizuka, the work by (Mcquiggan and Lester, 2007) also applies empathic behaviour in pedagogical agents. In the latter case, the authors propose a "data-driven affective architecture and methodology for learning models of empathy" (CARE). Their goal was to create and use a model of empathy within CARE, to respond accordingly in social situations. Creating the empathy model involves the realization of training sessions to register and associate the tasks performed by the agent to the corresponding empathic affective state of the companion's agent. Then, the induced model is ready to be used at run time. In this mode, the system continuously monitors the tasks that are performed done by the synthetic character controlled by the user. When it finds the same contextual conditions as the conditions simulated in the training sessions, the empathy model is used to elicit the corresponding empathic behavior in the companion character.

Although following an empirical approach does not restrict us to one specific theory, it does have however the drawback of being domain dependent while, at the same time, time consuming for data gathering in the training phase. Nevertheless, the achieved results suggest that it is well suited for devising new empathy models that, when applied to synthetic characters, allow them to display realistic empathic behavior. In fact, we believe that, in the future, new approaches that merge a theoretically based approach with an empirically based approach will lead to new and promising developments.

Another important area of application for computational empathy is in the creation of virtual companions for counseling. In that area, Bickmore (Bickmore, 2003) uses an embodied anthropomorphic animated character, named Laura, to study long-term social-emotional relationships, between virtual agents and humans. Laura's messages are triggered by a set of preconfigured conditions of the situation and are expressed through written dialogues and through a set of nonverbal cues. Although Laura did deliver empathic messages, it also had other relationship-building strategies in order to create this strong relation with the user. The results obtained suggest that relational agents have a positive impact on the user's perceived relationship with the agent, which "can build trusting, caring relationships with people that can be sustained over multiple interactions." In contrast with the aforementioned work, the empathic responses generated from our proposed model that are used for a virtual agent result from a general appraisal process, which allows a greater emergence of empathic emotions.

In addition to the focus on creating pedagogical agents that can express empathy, there is also research that focuses more on the study of pedagogical agents that are able to elicit empathic responses from the user. A well-known example in this area is the work of Paiva et al.(Paiva et al., 2005), a pedagogical system that addresses the bullying problem in schools. The system uses the FAtiMA agent architecture to model emotional behavior in the agents. There, the FAtiMA architecture was used to build concrete virtual characters featuring bullying situations, and it was tested with children to see how they responded to the actions of the characters, to teach them how to deal with bullying problems. This work goes further into using some theoretical elements of empathy, such as the use of the idea of proximity ("how close the learner will feel with the synthetic characters developed, in terms of the situation, behavior or even physical appearance") (Paiva et al., 2005), as an enhancement factor for empathic relations.

However, empathy has some cultural conditioning factors. In a more recent work on the FAtiMA architecture, an extension to capture culturally specific behavior was built, and was subsequently used to create a serious game (ORIENT) to foster intercultural empathy (Aylett and Paiva, 2012) (Mascarenhas et al., 2010). Note that although the appraisal, memory, planning and coping mechanisms of FAtiMA were extensively used in those applications (FearNot! and ORIENT), the agents themselves did not respond in an empathic way toward the user or toward other characters. In other words, their actions were triggered by the situations and events that were happening in the world, but not by the emotions of the other characters. Such type of social responses to the emotions of others was not captured in the original FAtiMA architecture.

In contrast, in our work, our architecture has the capability of "simulating" the situations of others, very much like placing oneself in the shoes of the other character to predict what emotion she (or he) would experience. The work reported here goes further than the original FAtiMA architecture with such a capability, allowing for empathic agents to be created.

However, as argued by Gratch (Gratch, 2012) in his comment to Aylett and Paiva work (Aylett and Paiva, 2012), theories and models that "separate the process of appraisal from the other cognitive processes complicate unnecessarily the model." He also argues that we should strive to find the minimally required processes to achieve the desired functions. We very much agree with this argument, and the present model can be seen as a first attempt to look at those minimal processes that are associated with empathy by seeking inspiration from a model that originated with work conducted with non-human primates (Preston and de Waal, 2002).

In addition to virtual agents, research in the area of social robotics also addresses models of empathic behavior (Tapus et al., 2007) (Kozima et al., 2004). In a recent example with a chess playing companion (see (Pereira et al., 2011) and (Leite et al., 2012)), empathic behavior was implemented in a companion robot, the iCat. In that scenario, the robot observes two users playing chess and comments their moves during the game. Toward one of the players, its "companion", the robot behaves in an empathic manner, and toward the other, it behaves in a neutral way. Empathy is expressed mainly by specific empathic utterances (e.g., "you're doing great, carry on") and by facial expressions that mirror the inferred affective state of the companion. The robot infers the affective state by using its own appraisal mechanism to appraise the companion's game situation. To reinforce the empathic behavior, other mechanisms were implemented, such as doubling the frequency with which the robot directs its eye gaze toward the companion compared with toward the other player. In the study that was conducted, the results suggest that empathic companion robots can be perceived by their human mates as friendlier than those that do not show an empathic behavior.

Finally, one important question that must be considered is how to evaluate empathic agents, and what is the impact that they have in users. As argued by Aylett and Paiva with regard to this type of evaluation (Aylett and Paiva, 2012), there is no straightforward way to test a computational model such as this one in isolation from a concrete application and scenario. As such, one method is to assess the psychological impact that empathic agents have on users in a concrete scenario and application. Following that approach, Brave, Nass and Hutchinson (Brave et al., 2005) have conducted one of the most thorough studies in this area. In their research, the user and a virtual agent play blackjack against a disembodied dealer agent. After the game has been played, the users were asked how they felt while playing and their opinions about the agent. In this study, the agent's empathic behavior (empathic emotion expression and emphatic dialogue) is based on a set of pre-scripted rules that are triggered on single occasions. Nevertheless, the authors were able to demonstrate that empathic virtual agents can improve their social interaction experience with users, more than those that are not empathic, or even those that display only self-oriented emotions. Empathic agents were seen as being more caring, likeable, trustworthy and submissive than non-empathic agents, while no significant difference was found for intelligence and dominance attributes.

To sum up, we believe that the aforementioned findings show the importance and relevance of this area, in particular in creating empathic virtual agents. However, as illustrated by the previous examples, most of the research conducted so far has not been on creating a general model of empathy for societies of agents that is grounded on the internal processes that are involved in human empathy. In contrast, in our proposed model, those processes are accounted for in the agent's empathy.

8. Conclusion and Future Research

Virtual agents are becoming more and more popular in all types of application areas, such as education, health, military, and gaming, among others. This spread demands an improved human-machine and machine-machine social interaction. In response to this request, the scientific affective computing community is researching new ways of enhancing social interaction in virtual agents. Some attention has been given to empathy because it is seen as a strong characteristic that is responsible for pro-social behavior between humans, who are the primal font of inspiration when developing virtual agents.

With this vision, we have taken an analytical approach by reviewing the literature and attempting to identify the process that lies behind empathy and its main elements, to design and build a generic computational model of empathy. The resulting model was inspired by a perception-action paradigm and was then integrated into an affective agent architecture so that it could be used within an interactive virtual agent environment. As previously detailed, the model follows a strict theoretic notion of empathy, which is usually known as affective empathy. This view does not address other related concepts such as sympathy, emotional contagion, personal distress or theory of mind. Nevertheless, the addressed empathy view and its underlying emotional process enables an empathic outcome through emotional sharing.

Using our model of empathy, synthetic characters can perceive emotional cues and elicit empathic emotions, which emerge by the modulation of various factors (similarity, affective link, mood and personality). These emotions can then trigger empathic actions. To analyze how users respond to the empathic behaviors generated by the model, an experiment was conducted in which we created a small scenario and derived two conditions from it: in one condition, we used the empathy model, and in the other condition, we did not (the control condition). The idea of the experiment was to verify whether the empathy model produced an impact on the perceived qualities of the characters and whether that impact followed the expected results from a previous study (Brave et al., 2005). To evaluate these assumptions, we performed an evaluation with a group of 77 participants.

The results obtained show that the character that has a similar empathic response to the user, with regard to the event underway, had significant effects on all of the analyzed qualities. This character was seen as being more caring, likeable, trustworthy, intelligent, dominant and less submissive. For the character that had a different (opposite) empathic response from the user's own empathic response, we could verify that the results were the opposite; significant effects were found for three of the analyzed qualities. The character was seen as being less caring, likeable and trustworthy. We could also verify that, when using the empathic model, the users' favorite character became the character with the strongest empathic response. The achieved results show that the characters with empathic behavior were perceived significantly different from those without it. These results follow, to great extent, the expected results.

As future work, we would like to explore this model in different scenarios and with richer situations, for example having the model applied in a scenario in which the virtual agents does not only interact between themselves but also with users at the same time. Additionally, we could study the model's contribution to generating better character believability or to establishing long-term relations with human users. Knowing that it is commonly accepted that empathy promotes pro-social behavior we would also like to study the contribution of the model in a scenario of task collaboration.

We foresee that other modulation factors, such as familiarity and past experience can be considered in the model, and some of the previously used modulating factors, such as similarity, can be further explored. We also plan to enhance the model with a cognitive variant of empathy that can address other related phenomena of empathy such as perspective-taking or helping behaviour. This possibility could broaden the scenarios in which where the model could be applied. For example, we think that, with an enhanced model of empathy as described, we could use the model to implement simulators for decision-making scenarios in which models that are solely utility based have failed to provide acceptable results in the past.

References

Arnold, M.B., Gasson, J.A., 1954. Feelings and emotions as dynamic factors in personality integration, in: Arnold, M.B., Gasson, J.A. (Eds.), The Human Person: An Approach to an Integral Theory of Personality. New York: Ronald, pp. 294–313.

Aylett, R., Paiva, A., 2012. Computational modelling of culture and affect. Emotion Review 4, 253–263.

Baron-Cohen, S., Wheelwright, S., 2004. The empathy quotient: an investigation of adults with asperger syndrome or high functioning autism, and normal sex differences. Journal of autism and developmental disorders 34, 163–175.

Bates, J., et al., 1994. The role of emotion in believable agents. Communications of the ACM 37, 122–125.

Batson, C., Shaw, L., 1991. Evidence for altruism: Toward a pluralism of prosocial motives. Psychological Inquiry 2, 107–122.

Bickmore, T.W., 2003. Relational agents : Effecting change through human-computer relationships. Ph.D. thesis. MIT.

Botvinick, M., Jha, A.P., Bylsma, L.M., Fabian, S.a., Solomon, P.E., Prkachin, K.M., 2005. Viewing facial expressions of pain engages cortical areas involved in the direct experience of pain. NeuroImage 25, 312–9.

Boukricha, H., Nguyen, N., Wachsmuth, I., 2011. Sharing emotions and space: Empathy as a basis for cooperative spatial interaction, in: Intelligent Virtual Agents, Springer. pp. 350–362.

Boukricha, H., Wachsmuth, I., 2011. Empathy-based emotional alignment for a virtual human: A three-step approach. KI-Künstliche Intelligenz 25, 195–204.

Boukricha, H., Wachsmuth, I., Carminati, M.N., Knoeferle, P., 2013. A computational model of empathy: Empirical evaluation, in: Proceedings of the Fifth Biannual Humaine Association Conference on Affective Computing and Intelligent Interaction (ACII), Institute of Electrical and Electronics Engineers (IEEE). pp. 1–6.

Brave, S., Nass, C., Hutchinson, K., 2005. Computers that care: investigating the effects of orientation of emotion exhibited by an embodied computer agent. International Journal of Human-Computer Studies 62, 161–178.

Campbell, M.W., de Waal, F.B.M., 2011. Ingroup-outgroup bias in contagious yawning by chimpanzees supports link to empathy. PLoS ONE 6, e18283.

Davidson, R., 1994. On emotion, mood, and related affective constructs, in: Ekman, P., Davidson, R. (Eds.), The nature of emotion: Fundamental questions. Oxford University Press, pp. 51–55.

Davis, M.H., 1983. Measuring individual differences in empathy: Evidence for a multidimensional approach. Journal of personality and social psychology 44, 113–126.

Davis, M.H., 1994. Empathy: A social psychological approach. Westview Press.

Deshmukh, A., Castellano, G., Kappas, A., Barendregt, W., Nabais, F., Paiva, A., Ribeiro, T., Leite, I., Aylett, R., 2013. Towards empathic artificial tutors, in: Proceedings of the 8th ACM/IEEE international conference on Human-robot interaction, IEEE Press. pp. 113–114.

Dias, J., Paiva, A., 2005a. Feeling and reasoning: A computational model for emotional agents, in: Proceedings of 12th Portuguese Conference on Artificial Intelligence, EPIA 2005, Springer. pp. 127–140.

Dias, J., Paiva, A., 2005b. Feeling and reasoning: A computational model for emotional characters, in: Progress in Artificial Intelligence, Springer. pp. 127–140.

Doce, T., ao Dias, J., Prada, R., Paiva, A., 2010. Creating individual agents through personality traits, in: IVA'2010 - 10th International Conference on Intelligent Virtual Agents, Springer Berlin / Heidelberg, Philadelphia, USA. pp. 257–264.

Eisenberg, N., Miller, P., 1987. Empathy, sympathy, and altruism: empirical and conceptual links, in: Eisenberg, N., Strayer, J. (Eds.), Empathy and its development, Cambridge University Press. pp. 292–316.

Eisenberg, N., Morris, A.S., 2001. The origins and social significance of empathy-related responding . A review of empathy and moral development: Implications for caring and justice by M . L . Hoffman. Social Justice A Journal Of Crime Conflict And World Order 14, 95–120.

Ekman, P., 1994. Moods, emotions, and traits, in: Ekman, P., Davidson, R. (Eds.), The nature of emotion: Fundamental questions. Oxford University Press, pp. 56–58.

Elliott, C.D., 1992. The Affective Reasoner : A process model of emotions in a multi-agent system. Phd. Nothwestern University.

Feshback, N.D., 1978. Studies of empathic behavior in children, in: Maher, B.A. (Ed.), Progress in experimental personality research, Academic Press, New York. pp. 1–47.

Goldsmith, H.H., 1994. Parsing the emotional domain from a developmental perspective., in: Ekman, P., Davidson, R. (Eds.), The nature of emotion: Fundamental questions. Oxford University Press, pp. 68–73.

Gratch, J., 2012. The sciences of the artificial emotions: Comment on aylett and paiva. Emotion Review 4, 266–268.

Hoffman, M., 2001. Empathy and moral development: Implications for caring and justice. Cambridge University Press.

Hoffman, M.L., 1987. The contribution of empathy to justice and moral judgement, in: Eisenberg, N., Strayer, J. (Eds.), Empathy and its development. Cambridge University Press, Cambridge. chapter 4, pp. 47–80.

John, O.P., Srivastava, S., 1999. The Big Five trait taxonomy: History, measurement, and theoretical perspectives, in: Pervin, L.A., John, O.P. (Eds.), Handbook of personality: Theory and research. second ed.. Guilford Press, New York, pp. 102–138.

Kozima, H., Nakagawa, C., Yano, H., 2004. Can a robot empathize with people? Artificial Life and Robotics 8, 83–88.

Krämer, N., Kopp, S., Becker-Asano, C., Sommer, N., 2013. Smile and the world will smile with you - the effects of a virtual agent's smile on users' evaluation and behavior. International Journal of Human-Computer Studies 71, 335–349.

Krebs, D.L., 1975. Empathy and altruism. Journal of Personality and Social Psychology 32, 1134–1146.

Latane, B., Darley, J.M., et al., 1968. Group inhibition of bystander intervention in emergencies. Journal of personality and social psychology 10, 215–221.

Leite, I., Pereira, A., Mascarenhas, S., Martinho, C., Prada, R., Paiva, A., 2012. The influence of empathy in human-robot relations. International Journal of Human-Computer Studies 71, 250–260.

Marsella, S., Gratch, J., Petta, P., 2010. Computational models of emotion, in: A blueprint for affective computing, pp. 21–45.

Marsella, S.C., Gratch, J., 2009. EMA : A process model of appraisal dynamics. Cognitive Systems Research 10, 70–90.

Mascarenhas, S., Dias, J., Prada, R., Paiva, A., 2010. A dimensional model for cultural behavior in virtual agents. Applied Artificial Intelligence 24, 552–574.

Mcquiggan, S., Lester, J., 2007. Modeling and evaluating empathy in embodied companion agents. International Journal of Human-Computer Studies 65, 348–360.

Ochs, M., Sabouret, N., Corruble, V., 2009. Simulation of the dynamics of nonplayer characters' emotions and social relations in games. IEEE Transactions on Computational Intelligence and AI in Games 1, 281–297.

Ochs, M., Sadek, D., Pelachaud, C., 2012. A formal model of emotions for an empathic rational dialog agent. Autonomous Agents and Multi-Agent Systems 24, 410–440.

Ortony, A., Clore, G., Collins, A., 1988. The cognitive structure of emotions. Cambridge University Press, Cambridge.

Paiva, A., Dias, J., Sobral, D., Aylett, R., Woods, S., Hall, L., Zoll, C., 2005. Learning by feeling: Evoking empathy with synthetic characters. Applied Artificial Intelligence 19, 235–266.

Paiva, A., ao Dias, J., Sobral, D., Aylett, R., Woods, S., Zoll, C., Hall, L., 2004. Caring for agents and agents that care: Building empathic relations with synthetic agents, in: AAMAS'2004 - International Conference on Autonomous Agents and Multi-agent Systems, ACM Press. pp. 194–201.

Pereira, A., Leite, I., Mascarenhas, S., Martinho, C., Paiva, A., 2011. Using empathy to improve human-robot relationships, in: Lamers, M., Verbeek, F. (Eds.), Human-Robot Personal Relationships. Springer Berlin Heidelberg. volume 59 of *Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering*, pp. 130–138.

Pizarro, D.A., Detweiler-Bedell, B., Bloom, P., 2006. Empathy, disgust, and moral persuasion, in: Kaufman, J.C., Baer, J. (Eds.), Creativity and reason in cognitive development. Cambridge University Press. chapter 4, pp. 81–98.

Plutchik, R., 1987. Evolutionary bases of empathy, in: Eisenberg, N., Strayer, J. (Eds.), Empathy and its development. Cambridge University Press, Cambridge. chapter 3, pp. 38–46.

Prendinger, H., Ishizuka, M., 2005. The empathic companion: A character-based interface that addresses users' affective states. Applied Artificial Intelligence 19, 267–285.

Preston, S.D., de Waal, F.B.M., 2002. Empathy: Its ultimate and proximate bases. The Behavioral and Brain Sciences 25, 1–20; discussion 20–71.

Rosenthal-von der Pütten, A.M., Krämer, N.C., Hoffmann, L., Sobieraj, S., Eimler, S.C., 2013. An experimental study on emotional reactions towards a robot. International Journal of Social Robotics 5, 17–34.

Rapson, R., Hatfield, E., Cacioppo, J., 1993. Emotional contagion. Cambridge university press.

Roseman, I., Smith, C.A., 2001. Appraisal theory: Overview, assumptions, varieties, controversies, in: Scherer, K., Schorr, A., Johnstone, T. (Eds.), Appraisal processes in emotion: Theory, methods, research. Oxford University Press, pp. 3–19.

Saarela, M.V., Hlushchuk, Y., Williams, A.C.d.C., Schürmann, M., Kalso, E., Hari, R., 2007. The compassionate brain: humans detect intensity of pain from another's face. Cerebral Cortex 17, 230–237.

Singer, T., Seymour, B., Doherty, J.O., Kaube, H., Dolan, R.J., Frith, C.D., 2004. Empathy for pain involves the affective but not sensory components of pain. Science 303, 1157–1162.

Stotland, E., 1969. Exploratory investigations of empathy, in: Berkowitz, L. (Ed.), Advances in experimental social psychology, Volume 4. Academic Press, New York. volume 4, pp. 271–314.

Tapus, A., Mataric, M.J., Scassellati, B., 2007. Socially assistive robotics. IEEE Robotics and Automation Magazine 14, 35.

de Vignemont, F., 2006. When do we empathize?, in: Fritch, C. (Ed.), Empathy and fairness. John Wiley & Sons, pp. 180–195.

de Vignemont, F., Singer, T., 2006. The empathic brain: how, when and why? Trends in cognitive sciences 10, 435–41.

de Waal, F.B.M., 2007. The 'Russian doll' model of empathy and imitation. Advances in Consciousness Research 68, 49–72.

Wagner, J., Kim, J., André, E., 2005. From physiological signals to emotions: Implementing and comparing selected methods for feature extraction and classification, in: IEEE International Conference on Multimedia and Expo, 2005. ICME 2005, IEEE. pp. 940–943.

Wicker, B., Keysers, C., Plailly, J., Royet, J.p., Gallese, V., Rizzolatti, G., Deusinglaan, A., Gerland, U.C.b.L., Garnier, A.T., 2003. Both of us disgusted in my insula : The common neural basis of seeing and feeling disgust. Neuron 40, 655–664.