Bio-Inspired Virtual Populations: Adaptive Behavior with Affective Feedback

Rui Filipe Antunes
MIRALab, Université de Genève
BioISI, Faculdade Ciências, Universidade Lisboa
(+41)223790109
rfantunes@fc.ul.pt

Nadia Magnenat-Thalmann
MIRALab, Université de Genève
(+41)223790969
thalmann@miralab.ch

ABSTRACT
In this paper, we describe an agency model for generative populations of humanoid characters, based upon temporal variation of affective states. We have built on an existing agent framework from Sequeira *et al.* [17], and adapted it to be susceptible to temperamental and emotive states in the context of cooperative and non-cooperative interactions based on trading activity. More specifically, this model operates within two existing frameworks: *a)* intrinsically motivated reinforcement learning, structured upon affective appraisals in the relationship of the agents with their environment [19,17]; *b)* a multi-temporal representation of individual psychology, common in the field of affective computing, structuring individual psychology as a tripartite relationship: emotions-moods-personality [7,15]. Results show a populations of agents that express their individuality and autonomy with a high level of heterogeneous and spontaneous behaviors, while simultaneously adapting and overcoming their perceptual limitations.

CCS Concepts
Computing methodologies→ Modeling and Simulation→Simulation Types and Techniques→ Artificial Life • Computing Methodologies→ Computer Graphics→Animation → Procedural Animation • Computing methodologies→Modeling and Simulation→ Simulation Types and Techniques→ Agent/Discrete models • Multi-Agent Systems.

Keywords
Crowd Simulation; Artificial Societies; Emotional Characters; Multi-Agent Frameworks; Artificial Life

1. INTRODUCTION
Generative populations of virtual humans, interacting autonomously and in real time, are an expanding area of research with immediate applications in domains such as urban simulations and games. The credibility of 3D scenes with autonomous populations of humanoid characters relies to a great extent, on their ability to show plausible behaviors. In that sense, the actions of individuals within groups of biological humans are not homogeneous nor periodic, and consequently, the heterogeneity and spontaneity of the behaviors performed by procedurally generated inhabitants are amongst the key factors for the quality and realism of the simulations.

CASA 2016, May 23-25, 2016,
ACM 978-1-4503-4745-7/16/05.
http://dx.doi.org/10.1145/2915926.2915929

We are interested in investigating how personality and emotionally influenced behaviors impacts actions, decisions and the expression of individual and global behaviors. We want to look at the trajectories of individual behavior as well as the patterns emerging at a global scale of the simulation resulting from the accumulation of individual life experiences. The objective is to be able to generate populations of agents that express their individuality and autonomy while simultaneously they adapt and overcome their perceptual limitations.

The paper is organized as follows, first we discuss the objectives and contextualize this work with related work in crowds and group simulation. Then, in Section 3, we provide details of the agency model. In section 4, we introduce a proof-of-concept, implementing this model in a population of autonomous trading agents (Fig. 1), and then we discuss the results of this experiment, bringing up the advantages and disadvantages of this model. Finally, we conclude suggesting future possibilities for research and development drawing on this approach.

1.1 Contributions
In this paper, we focus on one-to-one interactions within groups, and in particular, on the role played by psychology acting as part of an adaptive mechanism in shaping *i)* individual expressions and *ii)* group and crowd dynamics. We suggest a model of agency for generative systems that improves the level of individual spontaneity and expressivity while simultaneously permitting the progressive adaptation of agents to their environment.

Figure 1 General view of the generative population in the experimental setting, with two interacting individuals in the center of the image.

2. STATE OF THE ART

Attention to aspects of human psychology has been, traditionally, at the core of crowd research. Due to the relevant role played by emergency simulations in the development of the field, variables such as ‘level of panic', for instance, have been critical, strongly impacting the behavior of the virtual characters (Braun *et al.* [1], Tsai *et al.* [22]). The complexity of psychological traits varies from these simple ‘stress' emotional states (Park *et al.* [14]) to models where individuals differ from each other in their personality (Li *et al.* [11]). This is usually represented by means of multidimensional vectors, where each dimension corresponds to a personality feature. In Durupinar and colleagues [5], for instance, an agents' psychology is represented with the vector $v \leftarrow \{x, y, z, k, w\}$, standing for the five parameters of the OCEAN model (described in dimensions of openness, consciousness, extroversion, aggressiveness, and negativity) [3]. Similarly, Guy and colleagues use Eysenk's Three Factor [6], PEN (psychoticism, extroversion, and neuroticism). Behaviors of the virtual characters take these personality properties into account. One example is Bogdanovich and colleagues [21], where objects have annotated emotional responses. In the course of its actions, the agent chooses the object that better fits its personality parameters. With this type of approach, depending on the particular configurations, certain aspects of one individual might differ wildly from those observed in others, for instance, the level of tolerated proximity, the walking speed, etc.

In spite of these strategies to integrate human psychology in the simulations, the workings of personality and emotional states in humans is a complex phenomenon, dynamic, involving both internal and external processes. One-off short-termed emotions are inextricably linked with longer lasting moods and steadier temperaments, as well as actions performed by the individual and the environment. Modeling only stress levels or a rigid monolithic psychology, seems to be to too limiting, and other works attempt to bridge these gaps. One example is Silverman and colleagues' PMFServ [18, 2], where emotions are represented as dichotomies, joy-distress, hope-fear and like-dislike, resulting from actions performed by the agent (or its group's members). These emotions, then, have a direct impact on ‘stress' levels which, in turn, are used to trigger further motivated actions. Nonetheless, little importance is given to individual personality, which is represented as a predefined set of subjective beliefs informing the selection of actions performed. In turn, Navarro and colleagues [13] in SEstar use dynamic ‘stressometers' that articulate with higher level psychological traits modeled top-down during the initialization stage. These are narrowly defined features such as courage, stamina, etc.

Taking this scenario into consideration, Durupinar and colleagues improved their framework to include emotions, mood, and personality. Internal emotional states are differentiated and rendered visible using five different faces and body postures. Personality defines low-level behavioral attributes such as speed, and pushing behavior. Additionally, they add empathy and emotional contagion to their model. The drawback, however, is the cognitive dependence of emotions in their system. As they recognize, there is a wide range of emotions, such as surprise, which are never triggered using the appraisal heuristics that they have followed (The 22-emotions, OCC model [15]) [5, 4].

It seems opportune to address now this topic of research, and namely to bring together the dynamics of one-to-one interactions with the different dimensions of psychology and integrate these with reinforcement learning, in a model of agency that can be implemented successfully in virtual populations.

2.1 Related work

Given the complexity of environments composed of multiple interacting individuals in this agency model we have brought together some existing frameworks:

a) Reinforcement learning with affective states

We found it useful to endow our agents with the ability to address sequential and conflictual decisions. Reinforcement learning is a known technique of machine learning, inspired by behaviorist psychology [20]. The process resembles the evolutionary and adaptive process from living organisms, characterized by a progressive learning about the agent's environment, where perceptions become mapped to actions by trial and error. Typically, the possible action-states and the environment are modeled using a Markov chain with probabilities defined for each state transition. When changing state, the agent receives a reward according to the action performed, adapting the relative weight of the associated transition. Sequeira and colleagues [17], proposed a model for reinforcement learning, which we found relevant to our goal. They have built on earlier work from Singh [19], rewarding actions performed in dimensions that are simultaneously: *i)* affective - expressed by four appraisal dimensions: novelty, motivation, valence and control; and *ii)* functional - corresponding to the fulfillment of the agent's basic needs, for instance, being satiated [17].

b) Emotional tripartite scheme

Our goal is to combine this mechanism of affective agency with a more structured psychological scheme. The aim is to be able to model the effects of short-term emotions and longer term temperament in individual behavior. A standard tripartite psychological scheme, defined as personality-mood-emotion, is recurrent in the field of affective computing (Kshisagar and Magnenat-Thalmann [10], Gebhard [7], Santos *et al.* [16]). In these works, a basal personality is defined. This approach provides rooted patterns of behavior, or behavioral tendencies, predominantly located in specific zones of the psychological space. Emotions, on the contrary, are short-termed and consist of reactions to events. These works also model mood as an independent trait of the individual. The independence of this dimension is relevant since it allows incorporating both the memory from previous emotions, as well as their impact on personality.

c) Representation of temperament and emotions

From the works referenced above, Gebhard and Santos *et al.*, use Mehrabian's PAD multidimensional space to represent the mood of the agent. PAD is a system of representation designed to capture the entire domain of emotional experiences. Mehrabian argues that we can define every emotional trait with a three-dimensional vector consisting of three dimensions: *i)* pleasure-displeasure (P), *ii)* arousal-non-arousal (A) and *iii)* dominance-submissiveness (D), expressed in the bipolar space [-1,1] [12]. Furthermore, Mehrabian distinguishes transitory from long lasting emotions, what he describes, as emotional states and emotional traits [12].

In our model, the psychology of an agent is a dynamic process structured upon three layers: *i)* short-termed emotions, which result from interactions of the agent; and a temperamental factor, combining *ii)* a long-termed mood, which is the memory of these emotions; combined with *iii)* a biological imprint, which is a genetically determined component. For convenience, we use PAD dimensions for all three layers.

Traditionally, we see personality described using models such as the Big-5 or Eysenk (Krishsagar and Thalmann [10]) and emotions with OCC (Gehbard [7]), whereas PAD is more commonly seen representing moods (Santos et al. [16]). However, PAD is robust enough to provide representations of states on all three levels. Mehrabian clearly advocates the use of the model's dimensions to represent any "emotion or affect" [12]. Additionally, he establishes a direct correlation between PAD and the Big-Five model (For instance, extroversion = 0.24.pleasure + 0.72.dominance). As a consequence, and to simplify processes, we use PAD as a single model to represent the three layers: emotion-mood-personality.

We found it pertinent to combine these two frameworks and bring together: *i)* a tripartite and multi-temporal psychological scheme (personality-mood-emotion) - where long-term temperamental traits articulate with short-termed emotions, with *ii)* an action-selection mechanism - where perceptions are progressively mapped to actions in an evolutionary process of reinforcement learning (Fig. 2).

d) Metabolism and reproduction as intrinsic motivational factors

To test and experiment this model we have generated a population of autonomous individuals, organized in hierarchies and emulating life processes, with metabolic and genetic components. Closely aligned with works from ALife [8], a relatively simple motivational layer was implemented with agents exchanging token units of energy and resources, motivated by their own survival and perpetuation of genetic patrimony. Each individual is identified by a DNA-like string, which functions as a blueprint defining its own features. These phenotypic properties include aspects such as initial psychological parametrization, or the specification of class this individual belongs to. Class distinction between individuals allows us to establish functional hierarchies organized around the trade of resources and reproduction.

Additionally, agents emulate a life cycle, including birth, death (by lack of energy) and reproduction. When individuals multiply, their progeny inherits the parents' genetic blueprints, which are combined using operators of mutation and cross-over. When individuals are born, they appear in the animation from a predefined location. Similarly, when they die, they move to a specific location before they are removed from the system. This allow us to have continuous fluctuations of population density as well as discontinuous diversity of functions and roles.

Another noteworthy aspect is the role of resources, which is central in this scheme. They have a dual function: firstly, they are required to produce energy; secondly, they are recycled by the agent to produce other types of resources. To generate energy, each of the agents needs to use his own accumulated resources. As a consequence, he permanently needs to search for trading partners who might be interested in his resources and can trade the required ones. This provides a mechanism for intrinsic motivation to act in the world.

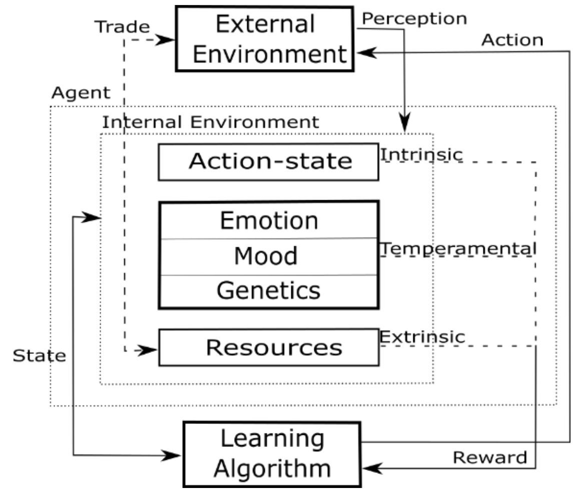

Figure 2 *Diagrammatic scheme of the model:* The agent makes an observation about the world (both the external environment, as well as its internal context). The affective reward of the action-state is influenced by its: *i)* functionality (extrinsic), *ii)* psychology (temperamental), and iii) the appraisal of the current state within the network (intrinsic). The success of interactions is dependent on the mood of the two agents involved and the outcome of the interaction generates a new emotion; adapted from Sequeira *et al.* [17].

3. MODEL DESCRIPTION

Emotions, mood and personality

The personality of each agent is established upon three interconnected layers: emotion, mood and personality. The one-off short-term emotions are inextricably linked with longer lasting moods and steadier temperaments as well as actions performed by individuals and their environment. We define each layer using PAD's nomenclature, which uses a three-dimensional vector consisting of: *i)* pleasure-displeasure (P), *ii)* arousal-non-arousal (A) and *iii)* dominance-submissiveness (D), expressed in the bipolar space [-1,1] [12].

To emulate biologic preconditioning, each agent is initialized with a static vector defining a dominant personality. Mood, on the contrary, is dynamic and composed of this biological imprint as well as the memory of previous emotions. Emotions are momentary events generated by: *a)* interactions with other agents and *b)* state-transitions in a Markov chain. When two individuals interact, the outcome of such an encounter is a function of: *a*) their respective moods, *b)* the utility of the encounter and *c)* the history of their previous encounters. An emotion is generated with an interaction, and the mood is updated accordingly. We express emotion as a function of appraisals in domains of motivation, valence, novelty, urgency and dominance. Emotion is represented as a PAD vector, where the first pair of appraisals determines the *Pleasure* dimension. The second pair defines *Arousal* and the last pair gives the value for *Dominance*. These appraisal factors are adapted from the agency model from Sequeira *et al.* to the context of cooperative and non-cooperative interactions based on emotional and temperamental states [13], which we will describe next (Eqs. 3.1 to 3.10).

Markov chain

The agent's behavior is established via a Markov chain. Transitions of state also generate emotions. Initially, all transitions leaving a state have identical weight. Each agent is constantly monitoring the internal and external environment, updating a set of sensors: *i)* lack of energy, *ii)* excessive libido, *iii)* proximity to an agent with resources, *iv)* proximity to a mate, *v)* neighbor is known, *vi)* currently in interaction, *vii)* emergency.

Once a change in one of these binary sensors is detected, a change of state is triggered. The agent then decides what action to choose based on his previous experiences. Once action *a* at state *s* is performed, the weight associated with the transitions of the action-state is then updated with reward *r* and a new emotion is generated. This reward is function of motivation, valence, novelty, urgency and dominance. Note however that when there is an emergency, the reward is hijacked since all the states have a transition to the emergency state with a predefined weight of value one.

3.1 Reward function

Sequeira [17] specifies r^{tot} resulting from an intrinsic reward r^{int}; that is calculated based on the appraisal equations that take in consideration the action-state, and an extrinsic reward r^{ext}, which results from external functional factors. By contrast, we combine affective, functional and intrinsic dimensions in each of the appraisal equations.

$$r^{tot}(s, a) \leftarrow \theta^m m(s, a) + \theta^v v(s, a) + \theta^n n(s, a) + \theta^u u(s, a) + \theta^c c(s, a) + \theta^d d(s, a),\ r : R \rightarrow [0, 1], \tag{3.1}$$

where θ^m, θ^v, θ^n, θ^u, θ^c and θ^d are scalar weight coefficients, and *m(s,a), v(s,a), n(s,a), u(s,a), c(s,a)* and *d(s,a)* the functions measuring the affect and functionality that results from performing action *a* at state *s*, which we will describe next.

3.2 Appraisal functions

Appraisals evaluate one-off momentary events and actions causing state transitions. As mentioned, all appraisals should take into consideration the temperament of the agent. However, we couldn't help but noticing parallels between these functions and the dimensions of pleasure, arousal and dominance, used in Mehrabian's model. As such, we found it useful to express each of the functions with its associated temperamental dimension.

3.2.1 Appraisal functions in the domain of pleasure

The level of pleasure resulting from performing action *a* at state *s*, is given by a combination of valence and motivation. Pleasure Valence is only measured in states involving an interaction and, consequently, in all other states only the appraisal of motivation contributes to pleasure.

a) Function of intrinsic *motivation* (eqs. 3.2.1 and eq. 3.2.2).

Departing from Sequeira's formulation of motivation as a function of the number of states required to reach the goal-state, we have defined motivation as the physical distance to goal *s**. However, when in the context of an interaction (distance is zero) trust is taken into account. For this purpose, the history of interactions with the partner is taken into account. *k* is the total of cooperative interaction with this partner and *m* the total of interactions. Furthermore, the influence of the pleasure component from the current mood is given by $\overrightarrow{M}^p$. α and β are weight coefficients for each of the factors. $\overrightarrow{M}^p$ is adjusted from its original bipolar PAD scale [1,1] to a normalized value [0,1].

$$m(s,a) = \alpha\frac{1}{1+d(s*)} + \beta\frac{\overrightarrow{M}^p+1}{2}, m: R \rightarrow [0,1], \tag{3.2.1}$$

$$m(s,a) = \alpha\frac{k}{m} + \beta\frac{\overrightarrow{M}^p+1}{2}, m: R \rightarrow [0,1], \tag{3.2.2}$$

where α, $\boldsymbol{\beta} \in [0, 1]$; $\overrightarrow{\mathbf{M}}^p \in [-1,1]$.

b) Extrinsic function of *valence* (eq. 3.3).

This measures the confidence of performing the action and it is given by the importance of this outcome *c* (cooperative: *c=1*; non-cooperative: *c=0*) in relation with the context of the previous experiences. This is calculated as the total number of cooperative interactions *w*, in relation with the total of interactions performed by the agent so far, *n*. As before, the influence of the component of pleasure in the current mood of the agent is given by $\overrightarrow{M}^p$. Similarly to the previous equation, α and β are weight coefficients,

$$v(s,a) = \alpha c\frac{w}{n} + \beta\frac{\overrightarrow{M}^p+1}{2}, v: R \rightarrow [0,1], \tag{3.3}$$

where α, $\boldsymbol{\beta} \in [0,1]$; $\overrightarrow{\mathbf{M}}^p \in [-1,1]$; $c \in \{0,1\}$.

3.1.2 Appraisal functions in the domain of arousal

Arousal or intensity of stimuli, is a function of the novelty of the situation, as well as the urgency of performing action *a*.

a) Function of *novelty* (eq. 3.4).

This function measures the degree of familiarity from the agent with the action-state. Sequeira defines $\lambda^{n_t(s,a)}$, as the number of times *t* the action *a* has happened at state *s*, where λ is constant, $\lambda < 1$. We have added a specific factor of novelty that is context dependent and is given by the number of previous interactions *q* that the agent *i* had with agent *j*. Because novelty is a function of arousal, the influence of the arousal component from the current mood is given by $\vec{M}^a$, α and β are weight coefficients,

$$n(s,a) = \alpha(\frac{\lambda^{n_t(s,a)} + \lambda^{n_q(i,j)}}{2}) + \beta\frac{\vec{M}^a + 1}{2}, n{:}\, R \to [0,1], \quad (3.4)$$

where α, $\boldsymbol{\beta} \in [0,1]$; $\vec{\mathbf{M}}^a \in [-1,1]$, $\lambda \in [0,1]$.

b) Extrinsic function of *urgency* (eq. 3.5).

This function integrates the agent's need for resources. The current level of resources is denoted by *r*, and *k* is an arbitrary minimum threshold, with λ as above. Equally, the influence of the arousal component from the current mood of the agent is given by $\vec{M}^a$. As before, α and β are weight coefficients.

$$u(s,a) = \alpha\lambda^{\frac{r}{k}} + \beta\frac{\vec{M}^a + 1}{2}, u{:}\, R \to [0,1], \quad (3.5)$$

where α, $\boldsymbol{\beta} \in [0,1]$; $\vec{\mathbf{M}}^a \in [-1,1]$, $\lambda \in [0,1]$.

3.2.1 Appraisal functions in the domain of dominance

This function models the familiarity of the agent with the present context, as well as the level of dominance-submission between two interacting individuals *i* and *j*. It is calculated in function of control and dominance.

a) Intrinsic function of *control* (eq. 3.6).

Control shows the experience of the agent, and this is expressed in function of the novelty of the present state. The influence of the dominance component from the current mood is given by $\vec{M}^d$. α and β are weight coefficients,

$$c(s,a) = \alpha\big(1 - \mathrm{n(s,a)}\big) + \beta\frac{\vec{M}^d + 1}{2}, c{:}\, R \to [0,1], \quad (3.6)$$

where α, $\beta \in [0,1]$; $\vec{M}^d \in [-1,1]$.

b) Extrinsic function of *domination* (eq. 3.7).

During interactions involving agents *i* and *j*, this equation assesses their power relation. It equates the ratios of successful interactions during their lifetimes $\frac{w}{n}$ and their need for resources (given by the urgency function - eq. 3.5). The existence of an interaction is signaled by the binary flag χ. *w* stands for the number of time the agent initiated a cooperative interaction, and *n* for the total of interactions. The component of dominance from the current mood of the agent is given by $\vec{M}^d$. Like before, α and β are weight coefficients,

$$d(s,a) = \alpha\chi\left(\frac{\left(\left(\frac{w}{n}\right)_i - \left(\frac{w}{n}\right)_j\right) + \left(u(s,a)_i - u(s,a)_j\right)}{2}\right) + \beta\frac{\vec{M}^d + 1}{2}, d{:}\, R \to [0,1], \quad (3.7)$$

where α, $\boldsymbol{\beta} \in [0,1]$; $\vec{\mathbf{M}}^d \in [-1,1]$, $\boldsymbol{\chi} \in \{0,1\}$.

3.3 Emotions

Emotions are translated directly from the appraisals described above. We define a vector of appraisals $\vec{A}$ using PAD dimensions, where pleasure component is $\vec{A}^p = (m(s,a) + v(s,a))/2$, arousal is $\vec{A}^a = (n(s,a) + u(s,a))/2$ and dominance $\vec{A}^d = (c(s,a) + d(s,a))/2$. However, we are using PAD's nomenclature to represent emotional traits where each dimension is a value in the space [-1,1]. Consequently, we need to convert each of the appraisals from their original scale [0,1]. In eq. 3.8, $\vec{E}$ is the PAD vector of emotion, $\vec{A}$ is the vector of appraisal rewards and $\vec{\mathrm{u}}$ is a unit vector.

$$\vec{E} = 2\vec{A} - \vec{\mathrm{u}} \quad, E{:}\, [0,1] \to [-1,1]. \quad (3.8)$$

3.4 Mood

The mood vector expresses the agent's basic personality as well as its memory of past emotions. Mood is rooted in the dominant personality, which is shaped by genetics (eq. 3.12) but it also evolves as a function of momentary emotions, resulting from state transitions and interactions (as described earlier).

To calculate the current mood, each of the different components of personality, genetics (G), emotion (E) and temperament (M) is taken into account. Consider the pleasure dimension, $\vec{M}^p(t)$, of individual *i* at time *t*. Its value is calculated in function of its genetic predisposition (given by $\vec{G}^p$) and the strength of the transitory emotional states (given by $\vec{E}^p(t)$) and the memory of past emotions ($\vec{M}^d{}_{(t-1)}$).

$$\vec{M}_{(t)} = \alpha\vec{M}_{(t-1)} + \beta\vec{E}_{(t)} + \overrightarrow{\gamma G}_{(t)}, M \in [-1,1], \quad (3.9)$$

where α, $\boldsymbol{\beta}$,$\gamma \in [0,1]$, $\vec{\mathbf{M}} \in [-1,1]$, $\vec{\mathbf{E}} \in [-1,1]$, $\vec{\mathbf{G}} \in [-1,1]$.

3.5 Genetic Preconditioning

In the moment of initialization, each agent receives a normalized PAD vector. Its parameters are constant values, which we will be referencing as $\vec{G}^p$,$\vec{G}^a$ and $\vec{G}^d$ accepting constant values *k1*, *k2*, and *k3*. This vector functions as an anchor, rooting temperament to an emotional tendency.

$$\vec{G} = (k1, k2, k3), G{:}\, R \to [-1,1]. \quad (3.10)$$

The outcome of this model is a multi-perspective description of the individual psychology. Dimensions of emotion, mood and genetic operate in multi-temporal scales generating expressive behaviors. Emotion traits can be easily integrated in animations to generate variations in individual behavior. For instance, the character may walk more or less exuberantly, and the speed of his movements may also vary accordingly. One aspect of great importance from individual psychology is expressed in one-to-one interrelationships, as we will discuss next.

3.6 Interactions

Agents act in the world in order to accomplish their motivated goals. Therefore, they interact with others and the course of these encounters is heavily influenced by their psychological states (Fig. 3).

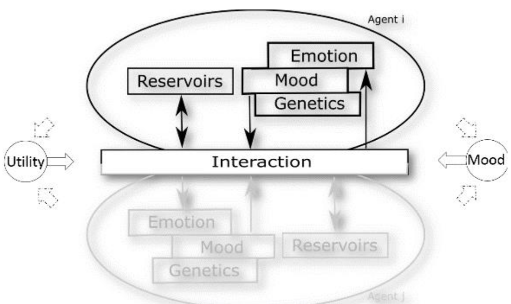

Figure 3 ***Diagram of an interaction:*** **To determine the success of the interaction, the mood and utility of the interaction for both agents is taken into consideration. The outcome will impact the emotions of both intervenients.**

As described earlier, in a feedback loop the outcome will also influence reciprocally the subsequent emotions of the agent. Equation 3.11 describes an interaction involving two agents *i* and *j*, determining the benefits for agent *i* of such an encounter. The function combines the psychology from both intervenients, $\psi_{i,j}(t)$; and an utilitarian evaluation, $\Upsilon_{i,j}(t)$, determining the usefulness of the interaction.

$$\textit{interaction} = \chi_{i,j}(t) = \psi_{i,j}(t) \cdot \Upsilon_{i,j}(t). \tag{3.11}$$

$$\textit{psychology} = \psi_{i,j} = \begin{cases} 1 \;\; if(\vec{M}_i(t) + \vec{M}_j(t))/2, Exuberant\,(+\mathrm{P} + \mathrm{A} + \mathrm{D}) \\ 0.25 \;\; if\left(\vec{M}_i(t) + \vec{M}_j(t)\right)/2, Bored(-\mathrm{P} - \mathrm{A} - \mathrm{D}) \\ 0.25 \;\; if(\vec{M}_i(t) + \vec{M}_j(t))/2\,, Dependent(+\mathrm{P} + \mathrm{A} - \mathrm{D}) \\ -1 \;\; if(\vec{M}_i(t) + \vec{M}_j(t))/2, Disdainful(+\mathrm{P} + \mathrm{A} - \mathrm{D}) \\ 1 \;\; if(\vec{M}_i(t) + \vec{M}_j(t))/2, Relaxed(+\mathrm{P} - \mathrm{A} + \mathrm{D}) \\ 0.25 \;\; if(\vec{M}_i(t) + \vec{M}_j(t))/2\,, Anxious(-\mathrm{P} + \mathrm{A} - \mathrm{D}) \\ 1 \;\; if(\vec{M}_i(t) + \vec{M}_j(t))/2, Docile(+\mathrm{P} - \mathrm{A} - \mathrm{D}) \\ -1 \;\; if(\vec{M}_i(t) + \vec{M}_j(t))/2, Hostile(-\mathrm{P} + \mathrm{A} + \mathrm{D}) \end{cases} \tag{3.12}$$

As per the psychology function (eq. 3.12), we average the vectors of mood $\vec{M}$ (which are given by eq. 3.9) from both intervenients *(i,j)* in the interaction.

The resulting vector indicates the openness of this encounter to be a cooperative one. For instance, when the result is a combination of moods considered as 'hostile', the function will return *-1* to indicate a non-cooperative stance. Conversely, if both intervenient are 'docile' to each other, the function will return *+1* to signal a cooperative mood.

The utility function results from a balance between the resources required by agent *i*, *(r)*, and the resources provided by agent *j*, *(p)*. Consider three resources *a*, *b* and *c*,

$$\textit{utility} = \Upsilon_{i,j}(t) = 1 - \frac{1}{1+\sum(r.p)}, \tag{3.13}$$

where r∈{a,b,c}, p∈{a',b',c'}

In a feedback loop, the dimensions of the emotional state of the two interacting agents will be updated after the transaction, as described earlier.

Figure 4 The expressivity of the gestures used in the behavioral animations reveals the internal state of the agents. In the center of the image, two individuals interact exhuberantly, while at their left another agent walks by showing a depressed locomotion.

4. EXPERIMENT

We wanted to be able to model the influence of emotions and personality in the interactions of a population of autonomous agents. To build this community (Fig. 4), we have implemented a society of agents structured in hierarchical layers and organized in a way to promote interaction and resource exchange. This type of multi-agent communities are characterized by permanent changes of patterns and flow since, in order to satisfy their goals, individuals need to move in space to find interaction partners. We found useful to include a similar and relatively simple motivational layer, where agents need to exchange token units of energy and resources, motivated by their own survival and the perpetuation of their genetic. Moreover, agents are equipped with a psychological mechanism that includes emotions, mood and personality as described earlier in the previous section. Finally, action are performed based on a Markov chain implemented using the reward mechanism described above. We further described 13 states (*i)* rest, *ii)* move to prey; *iii)* move to mate; *iv)*wander; *v)*found mate; *vi)*found prey; *vii)*facing a known mate; *viii)*facing an unknown mate; *ix)*facing a known prey; *x)* facing an unknown prey; *xi)* eat; x*ii)* reproduce; *xiii)* emergency exit).

To initialize the network, all transitions leaving one state have identical probabilities which are later updated according to the life-experiences of the agent. In equation 3.1 parameters θ^m, θ^v, θ^n, θ^u, θ^c and θ^d each have a value of 0.16. Parameters α and β in equations 3.2 to 3.7 have the value 0.75 and 0.25 respectively. Finally, in eq. 3.9 the parameters α, and β have the value 0.30 and γ the value 0.40.

To further analyze quantitatively the simulation we set it running for one hour and we have captured portraits of the population at intervals of one minute. The next sub-section describes the results of this run.

A video of the system running is available at (https://www.youtube.com/watch?v=_W0KEz52Ksw).

4.1 Results

Results will be analyzed in quantitative terms, in aspects such as: *a)* global heterogeneity of behaviors (Fig. 5); *b)* emotional dimensions (Figs. 6 and 7); and *c)* animation parameters (Fig. 8).

4.1.1 Actions performed

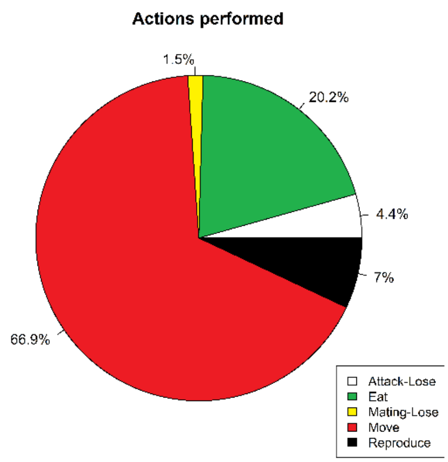

Figure 5 Activities performed by the agents throughout a run.

We started by looking throughout the run at the action being performed by each of the individuals at the moment of the snapshot. Fig. 5 shows that, as expected, agents spent most of their time (67%) walking. The rest of the time, they were involved in interactions. 25% of the time was spent in trading activities (Attack-Lose, Eat), and 9% mating (Mating-Lose, Reproduce). 27% of the time agents were successfully cooperating and only 5.9% they were found interacting with a non-cooperative attitude. The relatively high value of time spent interacting is justified by the limited number of goals that we have specified initially. This heterogeneity of actions results from the intrinsic need of the agents to capture energy and reproduce to satisfy their metabolic needs.

4.1.2 Emotional parameters

Fig. 6 depicts the psychological space of the population at an arbitrary moment of time. The genetic predisposition (personality) shows a uniform distribution with an occupation of the whole spectrum. This distribution was the expected result taking considering that a stochastic process initialized the initial population. Emotions and mood are less disperse and appear to be correlated. A large area of the spectrum appear occupied, but there seems to be some limitation with psychological states that are simultaneously positively aroused and positively dominant as well as simultaneously negatively pleasant and negatively dominant. The justification for this lays on the interdependency of functions from different domains, such as the control (dominance) being dependent on novelty (arousal).

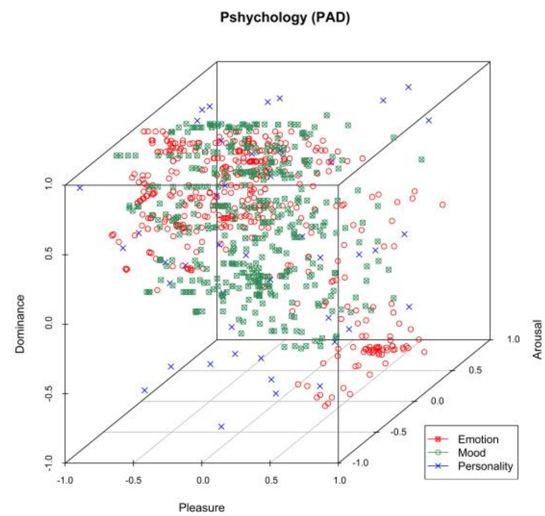

Figure 6 Graph with snapshot of the emotional domain of a population composed of 200 individuals, at an arbitrary moment in time.

We also have paid attention to individual differences between agents. Fig. 7 shows the sequence of psychological configurations of two arbitrary agents during a period of 15 minutes. As expected, and since their personalities differ they occupy different zones of the emotional spectrum. They also demonstrate a consistent continuity of emotional states. As discussed earlier, in our model personality impacts the mood and this is interdependent with emotions. The graph depicts this relationship with the agent's mood rooted on its personality. Similarly, the mood limits the freedom of its emotions. In turn, as defined by the model, moods are also affected by emotions and these two dimensions appear in the graph not only correlated but very close to one another.

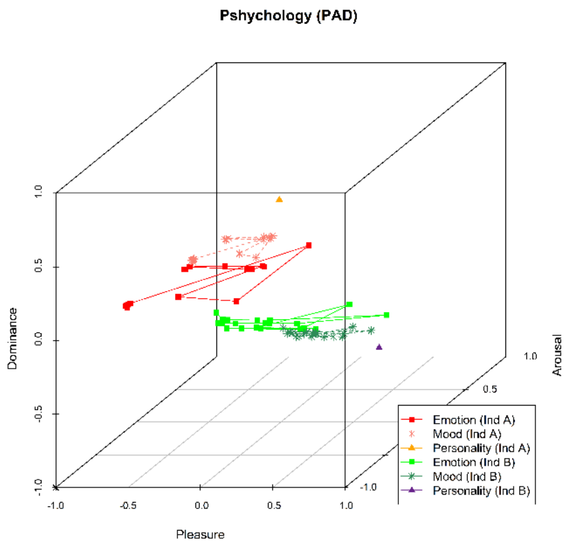

Figure 7 Evolution in time of the psychological state of two arbitrary individuals. Individual A (represented with reddish tones) occupies a zone from the psychological spectrum that is distinct from the one of Individual B (with greenish tones).

4.1.3 Animation parameters

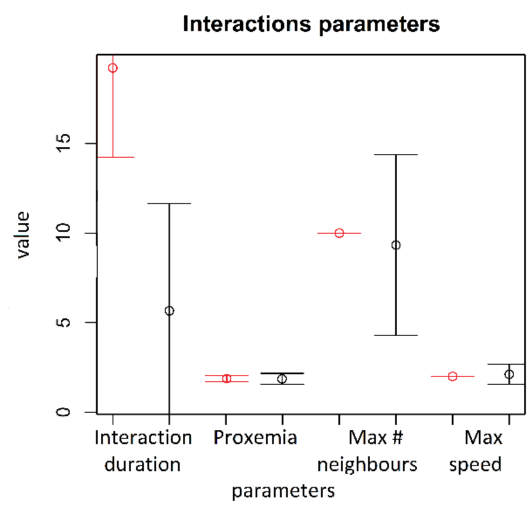

Figure 8 Parameters used in the visualization. From left to right, the duration of the interactions, the personal distance, the max number of neighbors, and maximum speed. In red, the run using a behavior tree (control), and in black, the agency model discussed in this paper (trial).

Finally, we have looked at the parameters of socialization used in the animations. We have contrasted this system with one without any emotional model and making use of a simple behavior tree instead of reinforcement learning (control) (Fig. 8). Results show a relative increase in our model (trial) in the rate of variation of the socialization parameters. The standard deviation of the duration of the interactions increased from 5.98 (control) to 8.42 (trial). This result is justified by the fact that in the control model, the duration of the interaction was subject to three binary factors: mating, resource exchange and cooperation, whereas in this model not only these three variables interfere with the duration of the interaction as this is also subject to the effects of the mood.

Significant variations were also found in the maximum number of neighbors, which was a constant value in the control model (control) and now (trial) had a deviation of 5.05.

Less dramatic variations were observed in the parameters of personal space and the maximum speed, since variation in these factors had to be restricted to a limit for the sake of the animation plausibility. Yet, we noticed an improvement on the degree of variability in both categories.

4.1.4 Walking towards a common goal

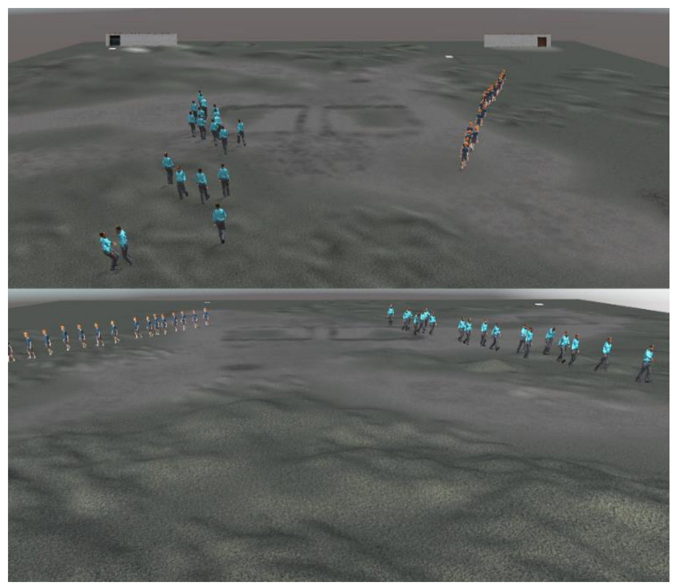

Figure 9 Contrast between our model of agency and one using RVO when individuals try to reach the common target.

Analyzing the model further, took us to look at how would it react to a standard benchmark such as walking towards a target. We have simulated a call to a classroom with two classes of individuals. The first group, dressed in dark blue, with a standard implementation of RVO, and another group dressed in light blue, with agents using our model. *Fig. 9-Top* shows the characters using RVO walking promptly and in line towards the classroom. The ones on the left, however, they linger around and take their time after the call to class is triggered walking slowly towards their goal.

Fig. 9-Bottom, shows the same two settings, only this time we have called the agents using our model with the sensor of urgency

activated. We can see that the response is different now, they walk disciplined towards the classroom. However, each one of the characters has its individuality, walking at his own speed and using different locomotion expressions and body postures.

4.1.5 Discussion

Our experiment shows a population acting heterogeneously with patterns of behavior commonly seen in ALife simulations: individuals behave spontaneously, with some aggregating and interacting in small groups of two or three while others walk at different speeds wandering through the terrain, and yet others walk but following one another. These behaviors are due to the underlying mechanism of agency used as the basis to build our system. Agents are intrinsically motivated to survive and perpetuate their genetic lineage, and for this purpose, they need to search permanently for partners of interaction. These basic principles from ALife form the basis to have an autonomous community always dynamic and active, exhibiting spontaneous and dynamic behaviors. Section 4.1.3 shows the level of heterogeneity of actions performed in a single run, from a limited set of 12 possible states defined in the Markov chain. This scenario seems to indicate a clear benefit of the integration of ALife's framework in this type of simulations, since it allows the relatively easy production of generative populations with self-motivated and autonomous individuals.

The levels of expressivity obtained are also noteworthy. In virtue of their dynamic psychological state, individuals in this community may differ in the way they perform identical behaviors. We achieve this diversity of expression with the inclusion of a tripartite model of personality in our model. We have implemented five different animations for each of the possible actions such as idle or walking or different types of behavior during the interactions. These variations are dependent on the current emotional state of the individual resulting in richer simulations showing an increased level of expressivity with emotional continuity.

Earlier, in order to improve the levels of expressivity of the simulations, Durupinar and colleagues have proposed an identical implementation of a model including emotions, moods, and personality. They similarly use five different animations for each behavior. However, their work builds upon OCC, PAD, and PEN models to define the psychological layers. In contrast, what is new about our approach is that we integrate the emotional pathways involved in decision making. Actions and emotional states become interdependent and operate in feedback loops. We use a single model (PAD) to represent different layers of psychology which allowed us to integrate affective dimensions with dynamic and continuous emotional states. Furthermore, coupling psychology with reinforcement learning resulted in the progressive adaptation of agents to their environment with affective-feedback. This feedback is dependent both on the outcome of the individual actions as well as on their psychological states. The learning process takes into account the emotional states generated by a multitude of factors such as the agents' experience, the distance, and complexity of steps to reach the goals, their need for energy and resources, their power relations, their confidence in others through their historical of betrayal and cooperation. The result of this interdependency is a mood that influences the agents' actions and swings smoothly with time as a consequence of its history of actions and reciprocally.

However, despite all the expressive richness of behavioral expression and heterogeneity, we also found some limitations in our model. As reported in section 4.1.2, we have noticed that agents cannot fully express all the potential range of emotions due to a design constraint where functions of different dimensions are interdependent.

5. CONCLUSIONS

Micro-scale behaviors are progressively more important in crowd research. Combining micro and macro dimensions allow us to easily change scales of simulation. For instance, in one moment we might have a bird's eye perspective from the roof of a building, and in the following, to be immersed in the multitude. The importance of expressive individuality becomes increasingly important in simulation contexts where realism is relevant.

We have looked at the role played by temporal variation in affective states in shaping individual behavior and expression. For this purpose, we have drawn on an existing agent-based framework [17], and adapted it to a context of cooperative and non-cooperative interactions based on trading behaviors. Furthermore, we have extended this framework to be susceptible to temperamental and emotional states of the agents. Agents' behaviors are defined using a Markov chain with dynamic probabilities updated using intrinsic reinforcement learning. Learning is consequent on appraisals of the autonomous interactions of the agent that are both functional and emotive.

We define a three-layered psychological model, integrating *i)* short-term emotions, which result from goal achievement; a temperamental factor, combining *ii)* a long-term mood, which is the accumulated memory of these emotions; and a *iii)* biological imprint, which is a genetically determined component of the personality of the agent. Mehrabian's PAD is used to represent these personality traits.

We further illustrate this model with its application in a generative virtual population composed of autonomous individuals that act heterogeneously, expressing rich and varied behaviors that are relatively consistent and coherent with their past actions. This population is composed of self-organizing social individuals that are able to adapt and prioritize their goals and behaviors. They are capable of autonomous and spontaneous interactions where personality and emotions play a relevant role. Moreover, the course of interactions is emotionally dependent and their quality is heavily dependent on psychological traits as they impact the interaction viability, outcome, and duration. Furthermore, the population density varies in time as new individuals are added while others are removed from the simulation, consequences of the agents' underlying trading activity.

In summary, this study is at its early stages, but it already offers promising results. We contribute an agency model for generative populations of humanoid characters based upon the variation of affective states. This model brings together reinforcement learning with individual emotions and personality. We show that when coupled with a framework inspired in ALife this model can bring quality to animations by enriching varied domains such as the heterogeneity of behaviors, their spontaneity and other parameters of interaction (duration, personal space, the maximum number of neighbors). The main advantage of such model is the relatively easy implementation of a self-organizing and autonomous community of virtual characters that are both a) rich in expression and b) intrinsically motivated to act in the world independently from the context.

This study is at its early stages, but it offers already promising results and some lines for future research. Next steps will include*:i*) adding different tasks and social roles in the simulations; *ii)* experiment different formulations for the appraisals, in particular limiting the currently existing interdependency, *iii)* adding more expressivity to the gestures and faces of the characters, with movements synthesized in real time; *iv)* incorporate reactive and cognitive layers of behavior. These steps will improve the realism and the complexity of behaviors.

6. ACKNOWLEDGMENTS

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 655226. The authors want to thank Joanna Pylak, Andra Chinchisan and Matthias Becker for helping revising the proofs, and the anonymous reviewers for their constructive comments.

7. REFERENCES

[1] A. Braun, B.J. Bodman, and S. R. Musse. Simulating Virtual Crowds in Emergency Situations. In Proceedings of the ACM Symposium on Virtual Reality software and Technology, 2005.

[2] D. Cassenti. Performance Moderated Functions Server (PMFserv) Military Utility: A Model and Discussion. Army Research Laboratory, 2009.

[3] P. Costa and R. McCrae. Revised NEO Personality Inventory (NEO PI-R) and Neo Five-Factor Inventory (NEO FFI). Psychological Assessment Resources, 1992.

[4] F. Durupinar, U. Gudukbay, A. Aman, and N. Badler. Psychological Parameters for Crowd Simulation: From Audiences to Mobs. IEEE Transactions on Visualization and Computer Graphics, 2015.

[5] F. Durupinar, N. Pelechano, J. Allbeck, U. Gudukbay, and N. Badler. How The Ocean Personality Model Affects the Perception of Crowds. IEEE Computer Graphics and Applications, 3:22-31, 2009.

[6] H. Eysenck and M. Eysenck. Personality and Individual Differences: A Natural Science Approach. Plenum Press New York, 1985.

[7] P. Gebhard. ALMA: A Layered Model of Affect. In Proceedings of the Fourth International Joint Conference on Autonomous Agents and Multiagent Systems. ACM. 2005.

[8] J. Holland. Hidden Order: How Adaptation Builds Complexity. Helix Books, 1996.

[9] H. Ji and G. Wakefield. Virtual World-Making in an Interactive Art Installation: Time of Doubles. In S. Bornhofen, J. Heudin, A. Lioret, and J. Torrel (eds.), Virtual Worlds: Artificial Ecosystems and Digital Art Exploration. Science ebooks, 2012.

[10] S. Kshirsagar and N. Magnenat-Thalmann. A Multilayer Personality Model. In Proceedings of 2nd ACM International Symposium on Smart Graphics, 2002.

[11] W. Li and J. Allbeck. Populations with Purpose. In Motion in Games, Vol 7060, LNCS, Springer, 2011.

[12] A. Mehrabian. Analysis of the Big-Five Personality Factors in Terms of the PAD Temperament Level. Australian Journal of Psychology, 48(2):86-92, 1996.

[13] L. Navarro, F. Flacher, and C. Meyer. SE-Star: A Large-Scale Human Behavior Simulation for Planning, Decision-Making and Training. In E. Bordini and Y. Weiss (eds.), Proceedings of the 14th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2015), pp 1939-1940, 2015.

[14] S. Park, C. Peng, F. Quek and Y. Cao. A Crowd Modeling Framework for Socially Plausible Animation Behaviors. In Motion in Games, pp 146-157, Springer, 2012.

[14] A. Ortony, G. Clore, and A. Collins. The Cognitive Structure of Emotions. Cambridge University Press, 1988.

[15] R. Santos, G. Marreiros, C. Ramos, J. Neves, and J. Bulas Cruz. Personality, Emotion and Mood Simulation in Decision Making. In New Trends in Artificial Intelligence. 14th Portuguese Conference on Artificial Intelligence, 2009.

[17] P. Sequeira, F. Melo, and A. Paiva. Emotion-Based Intrinsic Motivation for Reinforcement Learning Agents. In S. D'Mello, A. Graesser, B. Schuller, and J. Martin (eds), Affective Computing and Intelligent Interaction, Vol 6974, LNCS, pp 326-336. Springer, 2011.

[18] B. Silverman, M. Johns, J. Cornwell, and K. O'Brien. Human Behavior Models for Agents in Simulators and Game. Presence: Teleoperators and Virtual Environments, 15(2):163-185, 2006.

[19] S. Singh, T. Jaakkola, and M. I. Jordan. Learning Without State-Estimation in Partially Observable Markovian Decision Processes. In Proceedings of 11th International Conference on Machine Learning, pp 284-292. Morgan Kaufmann, 1994.

[20] R. S. Sutton and A. G. Barto. Introduction to Reinforcement Learning. MIT Press, 1998.

[21] T. Trescak, A. Bogdanovych, and S. Simoff. City of Uruk 3000 B.C: Using Genetic Algorithms, Dynamic Planning and Crowd Simulation to Re-enact Everyday Life of Ancient Sumerians. In Proceedings of the Simulation of the Past to Understand Human History Conference 2014.

[22] J. Tsai, N. Friedman, E. Bowring, S. Epstein, G. Kaminka, S. Marsella, A. Ogden, A. Rika, I. Sheel, M. E. Taylor, X. Wang, A. Zilka, and M. Tambe. Escapes: Evacuation Simulation with Children, Authorities, Parents, Emotions, and Social Comparison. In Proceedings of 10th International Conference on Autonomous Agents and Multiagent Systems, Vol 2, pages 457-464, 2011