Learning emotion regulation strategies: A cognitive agent model

Learning to cope with negative emotions is an important challenge, which has received considerable attention in domains like the military and law enforcement. Driven by the aim to develop better training in coping skills, this paper presents an adaptive computational model of emotion regulation strategies, which is inspired by recent neurological literature. The model can be used both to gain more insight in emotion regulation training itself and to develop intelligent virtual reality-based training environments. The behavior of the model is illustrated by a number of simulation experiments and by a mathematical analysis. In addition, a preliminary validation points out that it is able to approximate empirical data obtained from an experiment with human participants.

This chapter has been published as:
Bosse, T., Gerritsen, C., Man, J. de, & Treur, J. (2013). Learning Emotion Regulation Strategies: A Cognitive Agent Model. In *Web Intelligence (WI) and Intelligent Agent Technologies (IAT), 2013 IEEE/WIC/ACM International Joint Conferences on* (Vol. 2, pp. 245-252). IEEE.*

* Authors ordered alphabetically and contributed equally to this work.

7.1 Introduction

Emotions are commonly believed to play an important role in shaping human beings' behavior. They facilitate our decision making (Oatley & Johnson-Laird, 1987), prepare us for rapid motor responses (Frijda, 1986), and script our social behavior (Gross, 1998). Nevertheless, in case the valence and intensity of experienced emotions become sub-optimal, they can also hinder our performance. For example, when experienced stress and anxiety are high, different types of bias in decision making may occur, e.g., by creating a 'tunnel vision' (Klein, 1993). Fortunately, human beings are endowed with the ability to regulate their emotions, in order to achieve a more appropriate mode of functioning (e.g. Gross, 1998, 2001; Ochsner & Gross, 2005). In brief, *emotion regulation* refers to all (conscious or subconscious) processes humans undertake in order to affect their emotional response. Learning how to (better) regulate one's emotions is an important challenge, which has received considerable attention in various domains. For instance, police officers and military personnel take dedicated training in emotion regulation, to become more prepared to deal with extreme (negative) stimuli in the field (Berking et al., 2010).

The research reported here took place as part of a larger project that aims to develop a virtual reality-based training environment for professionals that often have to deal with negative stimuli. Application domains of such a system are public safety, public transport, and health care. The main idea of the system envisioned is that it generates a virtual scenario in which the trainee has to perform its task, while negative emotions are induced. Meanwhile the system monitors the behavior as well as some physiological aspects of the trainee. By combining these measurements with a computational model of emotion regulation, it will make an analysis of the development of the trainee's mental state with respect to experienced emotions. Based on this analysis, the system will provide appropriate feedback. For instance, in case it finds out that the trainee focuses too much on a particular negative stimulus, it will provide a cue to make the trainee aware of this.

As a first step towards the development of this system, this paper introduces an adaptive computational model of emotion regulation. In particular, the model distinguishes a number of emotion regulation strategies as put forward by Gross Gross (1998, 2001), and explicitly models the process of how persons learn to apply such strategies over time. In the next phase of the project, this model will be implemented in a virtual training environment, to enable the system to reason explicitly about the trainee's progress.

The remainder of this paper is structured as follows. In Section 7.2, an overview is given of the field of emotion regulation. Next, Section 7.3 describes the computational model, and Section 7.4 presents a number of resulting simulations for various parameter settings. Section 7.5 presents a mathematical analysis of the behavior of the model. Section 7.6 discusses some preliminary results of an experiment that has been performed to validate the model, and Section 7.7 concludes the paper with a discussion.

7.2 Background

Since a number of decades, emotion regulation has been an important subject of study in various research fields, varying from psychology and neuroscience to more technical disciplines such as AI. Whereas the literature in the former sheds more light on the psychological and neuro-physiological mechanisms behind emotion regulation, the latter has an emphasis on the development of computational models of emotion regulation. The following subsections briefly summarize the state-of-the-art in the respective areas.

7.2.1 Existing theories on emotion regulation

Probably the most influential research on emotion regulation of the last decades is the work by Gross (1998, 2001); Ochsner & Gross (2005). Gross (2001, p.215) defines emotion regulation as follows: 'Emotion regulation includes all of the conscious and nonconscious strategies we use to increase, maintain, or decrease one or more components of an emotional response'. The components considered are (1) the *experiential* component, (the subjective feeling of the emotion), (2) the *behavioral* component (behavioral responses), and (3) the *physiological* component (responses such as heart rate and respiration). Emotion regulation strategies cover 'antecedent-focused' regulation (in particular, *situation selection*, *situation modification*, *attentional deployment*, and *cognitive change*) and 'response-focused' regulation (in particular, *suppression* of a response). Note that situation selection and modification can be considered as more 'problem-focused' strategies, and the other ones as more 'emotion-focused'.

In addition to the emphasis on different strategies, various literature on emotion regulation addresses aspects like the neural bases of emotion regulation (Oatley & Johnson-Laird, 1987), the relation between emotion regulation (in particular fear extinction) and dreaming (Levin & Nielsen, 2007), and the effect of individual differences (Szczygieł et al., 2012).

7.2.2 Existing models of emotion regulation

Regarding computational models of emotion, a large body of literature is available in the field of affective computing (e.g. Becker-Asano & Wachsmuth, 2008; Bosse et al., 2010; Dias & Paiva, 2005; El-Nasr et al., 2000; Gebhard, 2005; Marsella & Gratch, 2009; Reisenzein, 2009; Soleimani & Kobti, 2012). The models addressed in that area can roughly be classified into two categories, namely those used to build in emotions into artifacts (such as virtual agents), and those used to assess emotional states of humans.[1]

One of the most influential models is EMA (Marsella & Gratch, 2009). This model takes appraisal theory as point of departure, which emphasizes that emotions are rooted in cognitions. Within EMA, intensities of emotional states are represented via real numbers in the [0..1] domain. Emotions arise when discrepancies between beliefs and desires (or other beliefs) are detected by automatic appraisal processes. Based on that perspective, a 'content model' is used, in which appraisal operates on rich symbolic representations of the emotion-evoking situation. Different types of emotion regulation (in particular down-regulation, also referred to as *coping*) are simulated as well by manipulation of these representations.

Despite a large number of successful applications, one aspect that has received less attention in EMA (and in most of the related approaches in the literature) is the *adaptivity* of emotion regulation capabilities over time. That is, agents are usually assumed to be equipped with a fixed number of regulation strategies, but the way in which these strategies can be learned or strengthened by training is under-represented. This is in contrast with the ability of human beings to become better (or worse) in emotion regulation based on practice (e.g. Berking et al., 2010). Hence, the current paper puts forward a novel computational model for emotion regulation, with a focus on adaptivity. The model is an extension of the model in Bosse et al. (2010).

7.3 Computational model

In this section the adaptive computational model for emotion regulation is described. First, an overview of model is given. In the Sections 3A-3C the different submodels are explained in more detail. The computational model presented here is based on mechanisms suggested in neurological literature. Recent neurological literature shows how emotion regulation takes place by an interaction between prefrontal cortex and amygdala (e.g. Kim et al., 2011; Phelps et al., 2004; Sotres-Bayon et al., 2004; Yoo et al., 2007). Several findings indicate that less adequate emotion regulation correlates to lower activity in prefrontal cortex areas and less strong connections from amygdala to prefrontal cortex (e.g. Kim et al., 2011). Moreover, strong indications have been found that REM-sleep strengthens both activation of prefrontal cortex and emotion regulation (e.g. Gujar et al., 2011).

The model is depicted in Figure 7.1. Here, the circles represent different states, which are all formalized in a numerical manner, in terms of a variable

[1] Nevertheless, many of the models developed can in principle be used for both purposes, like the model proposed in this paper.

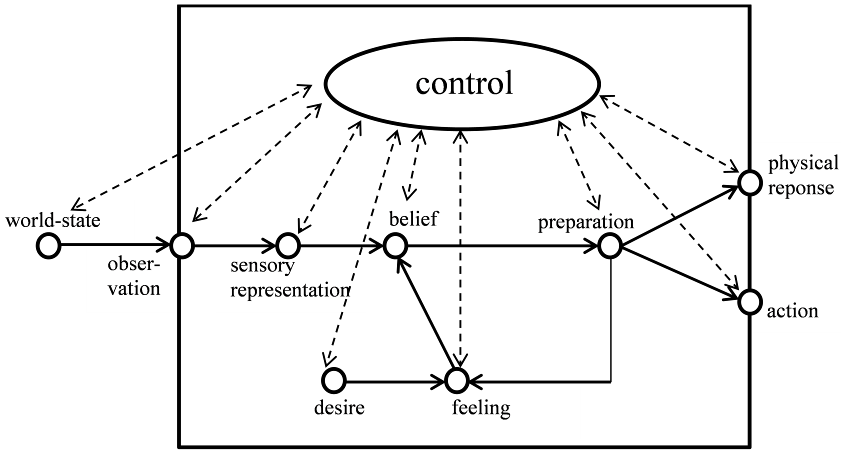

Figure 7.1: Overview of the Emotion Regulation Model

with a real value between 0 and 1. The influence of one state on another state is depicted by an arrow. The model that represents the emotion generation is depicted by using solid arrows. The *control* state represents the emotion regulation. Each state of the emotion generation model can be regulated by the *control* state. This is indicated by the dashed arrows. The *control* state has a suppressing effect on the other states. Further, all of the states have a positive effect on the control state, representing a kind of monitoring process.

7.3.1 Emotion generation

As a starting point, the model assumes a world state with a negative valence that is observed by the agent (e.g., a dead body). This leads to a sensory representation and a belief about this world state. Next, the agent prepares to act (e.g., to run away) and this preparation together with a desire (e.g., to avoid the stimulus) activate a feeling, which in turn may influence the belief state. This generation of feeling from preparation of emotional response follows the account based on an *as-if body loop* introduced by Damasio (1994). Following this theory, we use the term 'feeling' to refer to the mental state that emerges when an organism senses that its body is preparing to act. In contrast, Damasio reserves the term 'emotion' for the preparation state itself. In the model, we abstract from the specific 'type' of emotion (e.g., sadness, fear, anger) that is addressed, although we assume that it has a negative valence (we use 'fear' in our examples). Further, the preparation results in both a physical response (e.g., an increased heart rate) and an action.

Each of the states in the emotion generation model is generated by taking the value of that state at time point t and adding a fraction of the aggregated mean of values of states that have an influence minus the value of the state at time point t. For belief the formula would be as follows (where η represents the speed of activation spread):

$$q_{belief}(t + \Delta t) = q_{belief}(t) + \eta_{belief}[aggimpact_{belief}(t) - q_{belief}(t)]\Delta t$$

The aggregated mean in this example would be calculated by using the values of sensory representation, feeling and (in case of regulation, see next section) belief.control:

$$\begin{aligned} aggimpact_{belief} = {} & \omega_{sensoryrepresentation\_belief} * q_{sensoryrepresentation} \\ & + \omega_{feeling\_belief} * q_{feeling} \\ & + \omega_{belief\_belief} * q_{belief} \\ & + \omega_{bel.control\_belief} * q_{bel.control} \end{aligned}$$

The values are weighted by the connection strengths between the two states, represented by ω.

The generation of the other states is determined in a similar manner. Only for the state 'desire' no aggregated mean has been used since in the model this state is not influenced positively by other states.

7.3.2 Emotion regulation

To enable the agent to regulate the emotion levels for the different states in the model, control states have been added (depicted together as one oval in Figure 1). Each control state has a negative influence on the related state in the emotion generation model. To this end, the value of the downward connections from the control to the different states (i.e., the $\omega_{s.control\_s}$) is taken negative for all s. The upward connections (i.e., $\omega_{s\_s.control}$) are used to monitor the activation levels of the states. Their strength represents the extent to which the person is able to monitor (and regulate) that particular state. The emotion regulation strategies by Gross (2001) have been adopted:

- For *situation selection* and *situation modification* (see Section 7.2.1), the *world state* is altered (e.g., avoiding a stimulus).
- There are two variants of *attentional deployment* in the proposed model. First, the *observation* state is altered (e.g. by looking away from the stimulus). Secondly, in case of *sensory representation* the internal focus of attention is regulated (e.g., thinking about something else).

- *Cognitive change* is possible when the *belief* is regulated (e.g., reappraisal of the situation: saying to yourself that the situation is not bad). But it is also possible to regulate the *desire* by adjusting one's goals.

- The response-focused regulation strategy *suppression* is applied to *feeling* (e.g., suppressing feelings experienced), *physiological response* (e.g., showing a pokerface) and *preparation/action* (e.g., staying at a location instead of running away).

7.3.3 Adaptivity

Emotion regulation is assumed to be adaptive in the sense that by a learning process the upward connections to the control states can be strengthened gradually (see also the neurological justification at the start of this section). This adaptive process is modeled by a Hebbian learning principle (Hebb, 1949): strengthening of a connection over time may take place when both nodes are often active simultaneously ('neurons that fire together wire together'). The principle has recently gained interest both by more extensive empirical support (e.g. Bi & Poo, 2001), and more advanced mathematical formulations (e.g. Gerstner & Kistler, 2002). In the model this principle is applied to the upward connections to the control states (see Figure 7.1).

This principle is formalized as follows: for an upward connection from node i to node j, its strength ω_{ij} is adapted using the following Hebbian learning rule, taking into account a maximal connection strength 1, a learning rate η, and an extinction rate ζ:

$$\begin{aligned}\frac{d\omega_{ij}(t)}{dt} &= \eta q_i(t)q_j(t)(1-\omega_{ij}(t)) - \zeta\omega_{ij}(t)\\ &= \eta q_i(t)q_j(t) - (\eta q_i(t)q_j(t)+\zeta)\omega_{ij}(t)\end{aligned}$$

Here $q_i(t)$ and $q_j(t)$ are the activation levels of node i and j at time t and $\omega_{ij}(t)$ is the strength of the connection from node i to node j at time t. By the factor $1-\omega_{ij}(t)$ the learning rule keeps the level of $\omega_{ij}(t)$ bounded by 1 (which could be replaced by any other positive number); Hebbian learning without such a bound usually provides instability. When the extinction rate is relatively low, the upward changes during learning are proportional to both $q_i(t)$ and $q_j(t)$ and maximal learning takes place when both are 1. Whenever one of $q_i(t)$ and $q_j(t)$ is close to 0 extinction takes over, and ω_{ij} slowly decreases. In principle, the learning rate η and extinction rate ζ can be taken differently for the different connections. In the simulations discussed in Section 7.4, always the following values have been used: $\eta = 0.01$ and $\zeta = 0.001$.

7.4 Simulation results

To test the model, a number of simulation experiments have been performed. The results are depicted in Figures 7.2 to 7.4. In these figures, time is shown on the horizontal axis and the activation values of the states are shown on the vertical axis.

Figure 7.2 shows the simulation of the model without any regulation. Throughout this simulation, the activation value of the world state (q_{world}) and the desire (q_{desire}) have been set to 0.8 and 0.7, respectively. This corresponds to a case where the person is exposed to a rather extreme stimulus (q_{world}), while also having a fairly strong desire to avoid this (q_{desire}). All other activation states start with a value of 0. In this simulation the no regulation condition is realized by setting the $q_{s.control}$ and $\omega_{s\_control}$ parameters for all states s to 0. Moreover, all η_s are 0.1 (representing a low speed of activation spread), except for the action state ($\eta_{action} = 0.05$, otherwise that state would not be distinguishable in the graphical representation from the physical response). The values of all positive connection strengths that lead towards the same state always sum up to 1, whereas the values of all negative connection strengths (i.e., the downward connections $\omega_{control\_s}$ from the control state) have been set to -0.5 (except for the connection to the world state).

As shown in Figure 7.2, without regulation, the activation values of all states either approximate the *world state* (for *observation* and *sensory representation*) or a value between the activation values of the *world state* and the *desire*. In Figure 7.3 the results are shown for a situation in which all states are regulated. The parameters are equal to the parameters for the simulation in Figure 7.2, with the exception that the connection strengths from all states to the control (i.e., all $\omega_{s\_control}$) now have the value 0.5. This represents the case that the person already starts with some average regulation skills regarding all types of strategies. As shown in Figure 2b, these parameter settings result in a scenario where the activation of all states first increases, but is then suppressed because of the regulation. The closer a state is to the *world stat*e (see Figure 7.1), the higher the activation value in which it ends up.

Figure 7.4 shows a simulation of the model in which only the *belief* state is regulated extremely well. This corresponds to individuals that are very skilled at the strategy *reappraisal*. This was realized by taking $q_{belief.control} = 0.5$, $\omega_{control\_belief=-1}$, and $\omega_{control\_s=-0.1}$ for all other s. The figure clearly shows that the *belief* state, as well as all states that 'depend' on this state, end up at much lower values than in Figure 7.2, whilst the other states remain high.

Finally, Figures 7.5 and 7.6 show simulations of the model in which only the respective states *observation* and *feeling* are regulated. This corresponds to individuals that are very skilled at the strategies *attentional deployment* and *suppression*, respectively. For Figure 7.5, this was realized by taking

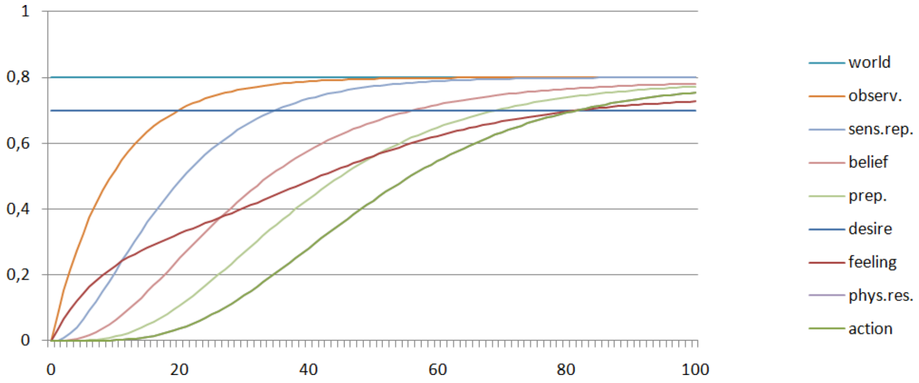

Figure 7.2: Simulation results without regulation

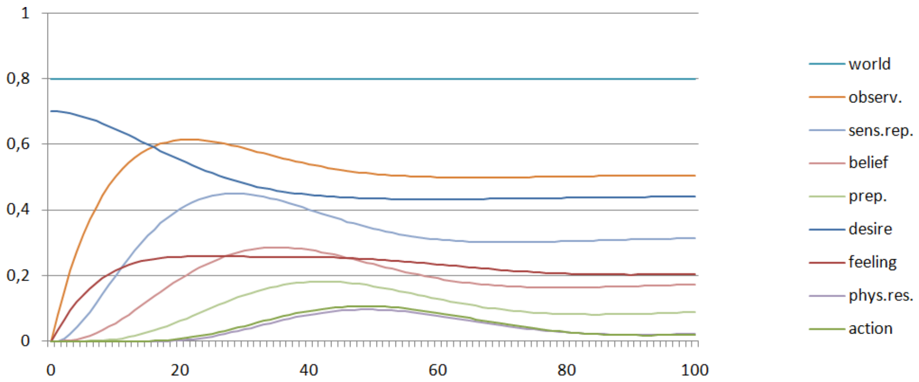

Figure 7.3: Simulation results with regulation

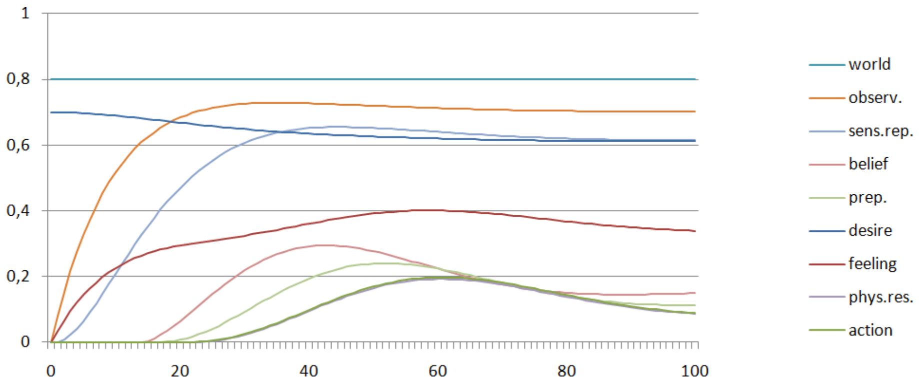

Figure 7.4: Simulation results with only *belief* regulated

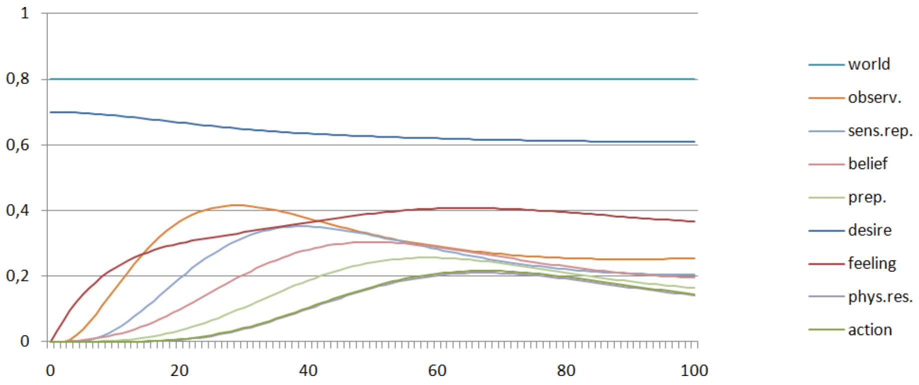

Figure 7.5: Simulation results with only *observation* regulated

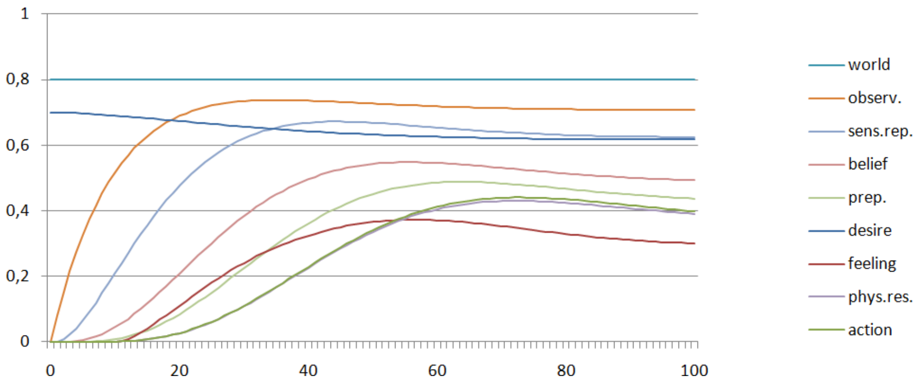

Figure 7.6: Simulation results with only *feeling* regulated

$q_{observation.control} = 0.5$, $\omega_{control\_observation} = -1$, and $\omega_{control\_s} = -0.1$ for all other s. For Figure 7.6, this was realized by taking $q_{feeling.control} = 0.5$, $\omega_{control\_feeling} = -1$, and $\omega_{control\_s} = -0.1$ for all other s.

Like in Figure 7.4, Figures 7.5 and 7.6 show that the state that is explicitly regulated (i.e., either the *observation* or the *feeling* state), as well as all states that 'depend' on this state, end up at much lower values than in Figure 7.2, whilst the other states remain high.

To illustrate the adaptive nature of the model simulations have been run as shown in Figures 7.7 and 7.8. Here, the simulation time has been extended from 100 to 400 time steps, to be able to simulate multiple learning trials. To

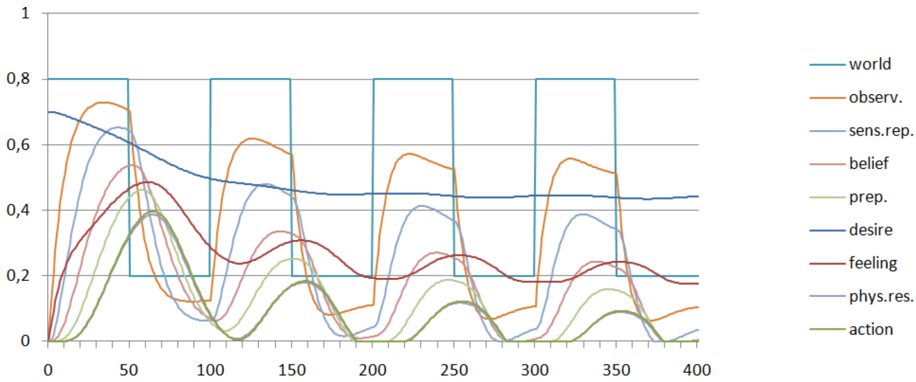

Figure 7.7: Learning effect over multiple trials

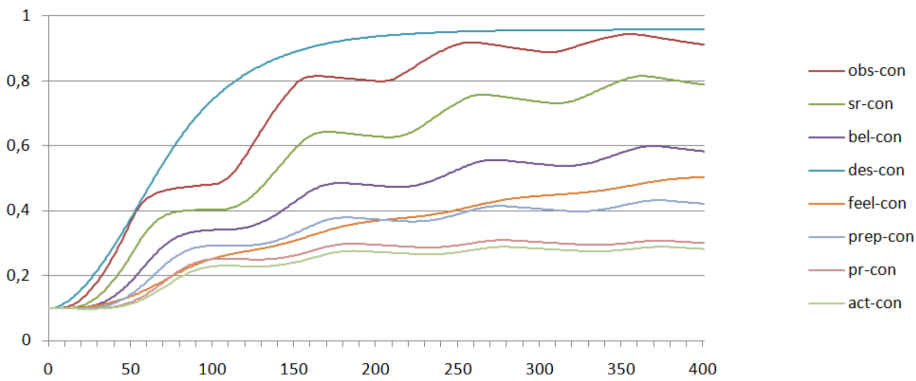

Figure 7.8: Connection strengths

this end, the *world state* has a value of 0.8 for 50 time steps, then a value of 0.2 for the next 50 time steps, and so on. Here, at the start the person is relatively bad at emotion regulation (all $\omega_{s\_control}$ start at 0.1, while all other parameter settings are equal to Figure 7.3), but this improves over time. Figure 7.8 shows that, while the person is exposed to the stimulus, all connections towards the control state are strengthened (following the Hebbian learning), whereas due to extinction they slowly decrease during the intervals that there is no strong stimulus. The consequence of this is that the person gradually improves her regulation skills over time, which results in the ability to bring the corresponding activations further down over the trials (as shown in Figure 7.7).

7.5 Mathematical analysis

Usually dynamic properties of agent models can be analyzed by conducting simulation experiments, as done in the previous section. In this section some mathematical methods are described to analyze such properties: equilibria, increasing/ decreasing trends, speed of convergence to equilibria.[2] An *equilibrium* is a state in which no change occurs. Based on the difference or differential equations describing the model, it can be analyzed which states are equilibria. Moreover, when a variable is not in equilibrium, it can be found when it is increasing or decreasing. Consider the dynamic model for Hebbian learning for the strength ω of a connection from a state $S1$ to a state $S2$ with maximal connection strength 1, learning rate $\eta > 0$, and extinction rate $\zeta \geq 0$ (here q_{S1} and q_{S2} denote the activation levels of the states $S1$ and $S2$):

$$\begin{aligned} \omega(t+\Delta t) &= \omega(t) + [\eta q_{S1}(t) q_{S2}(t)(1-\omega(t)) - \zeta\omega(t)]\Delta t \qquad (7.1)\\ \frac{d\omega}{dt} &= \eta q_{S1} q_{S2}(1-\omega) - \zeta\omega \end{aligned}$$

From these expressions, the following can be analyzed:

$$\text{Increasing:} \qquad \frac{d\omega}{dt} > 0 \Leftrightarrow \omega < \frac{\eta q_{S1} q_{S2}}{\zeta + \eta q_{S1} q_{S2}} \qquad (7.2)$$

$$\text{Equilibrium:} \qquad \frac{d\omega}{dt} = 0 \Leftrightarrow \omega = \frac{\eta q_{S1} q_{S2}}{\zeta + \eta q_{S1} q_{S2}} \qquad (7.3)$$

$$\text{Decreasing:} \qquad \frac{d\omega}{dt} < 0 \Leftrightarrow \omega > \frac{\eta q_{S1} q_{S2}}{\zeta + \eta q_{S1} q_{S2}} \qquad (7.4)$$

So, the value of ω increases when it is under the equilibrium value $\underline{\boldsymbol{\omega}}$ and decreases when it is above this value: it is an attracting equilibrium. This indeed is observed in the simulation experiments. Note that this equilibrium value $\underline{\boldsymbol{\omega}}$ usually is lower than 1. It may be close to 1, but when $\zeta > 0$ it never will be equal to 1. In fact the maximal value of this equilibrium is when both $q_{S1} = 1$ and $q_{S2} = 1$, in which case the equilibrium value $\underline{\boldsymbol{\omega}}$ is $1/(1+\zeta/\eta)$. For example, for $\eta = 0.4$, $\zeta = 0.08$, and $q_{S1} = 1$ and $q_{S2} = 1$, this is 0.83.

A useful approach to analyze the behavior close to equilibria is by determining a linear approximation of the system. This can be done by considering the distance δ_i of y_i to an equilibrium value $\underline{\boldsymbol{y_i}}$ by writing $y_i = \underline{\boldsymbol{y_i}} + \delta_i$ with $\underline{\boldsymbol{y_i}}$ the equilibrium value. This approach will be used for the case of Hebbian learning by writing $\omega = \underline{\boldsymbol{\omega}} + \delta$. For this analysis it is assumed that $q_{S1}q_{S2} = c$ for a constant c. Substituting $\omega = \underline{\boldsymbol{\omega}} + \delta$ in the differential equation for ω

[2]See: http://www.cs.vu.nl/~tbosse/STRESS/mathan.pdf (accessed 22-08-2013).

provides:

$$\frac{d(\underline{\boldsymbol{\omega}} + \delta)}{dt} = \eta c(1 - (\underline{\boldsymbol{\omega}} + \delta)) - \zeta(\underline{\boldsymbol{\omega}} + \delta) \tag{7.5}$$
$$\frac{d\delta}{dt} = \eta c(1 - \underline{\boldsymbol{\omega}}) - \zeta\underline{\boldsymbol{\omega}} - \eta c\delta - \zeta\delta$$

(as $\underline{\boldsymbol{\omega}}$ is an equilibrium value it holds $\eta c(1 - \underline{\boldsymbol{\omega}}) - \zeta\underline{\boldsymbol{\omega}} = 0$)

$$\frac{d\delta}{dt} = -\eta c\delta - \zeta\delta \tag{7.6}$$
$$\frac{d\delta}{dt} = -(\eta c + \zeta)\delta$$

This is a simple type of a single linear differential equation (where c is assumed constant) for which an analytic solution is known to be exponential:

$$\delta(t) = \delta(0)e^{-(\eta c+\zeta)t} \tag{7.7}$$

This shows that the attraction of equilibrium $\underline{\boldsymbol{\omega}}$ goes according to a negative exponential pattern with the given exponent $-(\eta c + \zeta)t$. In particular, every time unit the difference δ is multiplied by $e^{-(\eta c+\zeta)t}$. Note that for $c = 0$ (no activation), the exponent is $-\zeta t$ (and $\underline{\boldsymbol{\omega}} = 0$), which implies that, for a low extinction rate, the downward convergence to 0 is very slow, according to a convergence rate ζ. For $\zeta = 0.08$ this $e^{-\zeta}$ is 0.92. In contrast to this, for $c > 0$, the $e^{-(\eta c+\zeta)}$ is lower, depending on the value of c. For example, for $\eta = 0.4$, $\zeta = 0.08$, and $q_{S1} = 1$ and $q_{S2} = 1$, it is 0.62. The theoretically derived values for equilibria and convergence speed have been compared with the observed values in simulations as described in Section 7.4. The deviation between the observed and theoretical values found was less than 0.01.

7.6 Preliminary results

As a first step towards validation of the model, an attempt has been made to replicate some empirical findings gathered in an experiment to investigate the effects of using virtual training to reduce subjects' response to emotional stimuli. The following section describes this data in more detail. Afterwards, results obtained using the model described above are compared to a linear approximation of the empirical data, illustrating the model's capacity to mimic learning of emotion regulation as exhibited by human subjects.

7.6.1 Empirical data

The data used for this preliminary validation is obtained from the experiment described in Bosse et al. (2012).[3] Here, two groups of participants were asked to rate 150 images taken from the IAPS picture set on their emotional intensity on a scale from 1 to 5. This was done once in the morning and again six hours later in the afternoon. One group, the training group, returned exactly three hours after the first measurement to undergo a training where each picture was shown again and the participants were asked to apply reappraisal techniques to control their emotional response (i.e. imagining the image was digitally altered). In both groups results show an increase in the number of pictures that are given the lowest rating in the afternoon. However, for the higher ratings, the training group clearly shows a shift towards lower ratings, which is not visible in the control group. Moreover, the training group gave on average significantly lower ratings. This effect was even more pronounced when only considering all pictures with a valence smaller than 5 on the standard IAPS classification (the scale of IAPS valence is 1-9). The data for these 57 'negative' stimuli is used below.

7.6.2 Reproducing the data

The experiment was simulated by recreating the experimental setup as described in Bosse et al. (2012) and using the objective arousal for each picture as provided by the IAPS classification (the scale of IAPS arousal is 1-9) as input for the model. First though, the arousal needed to be linearly scaled to a value between 0 and 1. Participants were asked about the emotional intensity they experienced, hence the activation of 'feeling' in the model was selected as the output to be matched to the empirical data. Although it is difficult for people to express their emotional feeling accurately, for this first validation step it would suffice to take the maximal activation value of 'feeling' in response to each stimulus directly (scaled to a 1-5 value). The training session was simulated by presenting each stimulus to the model again, but now with an increased learning rate for belief control corresponding to the given task of cognitively changing the emotional impact of the stimulus by reappraising the content of the picture. For the control group, the model was run an equal length of time with zero input. Thereafter, similar as before, ratings for each image were gathered with the model, thereby simulating the afternoon measurement.

Initial parameters were manually tuned for the morning measurement and kept the same for both groups of participants. Furthermore, linear approximations were made for the morning data of both groups and re-used to approxi-

[3] Chapter 5 (p.89) describes an extension of that experiment.

mate the ratings given in the afternoon. These results are used as a comparative baseline for the proposed model.

Figures 7.9 and 7.10 show the error of both approximations compared to the empirical data. The stimuli were grouped in four sets based on their arousal and the results were averaged for these groups. The errors shown are the differences between these averages. As shown, the linear approximation gives more accurate results for the control group, except for stimuli with a medium arousal. However, for the training group, the model is more accurate, except for stimuli with low arousal in the morning measurement. Moreover, it can be clearly seen that the linear model has no method for approximating emotion regulation learning, causing large errors in the afternoon measurement.

If we take the average rating of each of the four sets of images both in the morning and in the afternoon, we can use the difference between them as a measurement for learning emotion regulation. This measurement can then be used as a tool to investigate the success of the model in approximating this learning by comparing its difference with the measurement found in the empirical data. Figure 7.11 shows for both the model and the linear approximation the error between both differences.

These graphs show that the linear approximation, where no learning occurs, performs poorly. More interestingly, the model proposed in the current paper produces smaller errors for both groups. Although this will not be the final word on how well this model simulates learning of emotion regulation, these first results indicate that it might be able to replicate empirical data such that both the learning of emotion regulation that takes place in the control group as well as the explicit learning during a training can be incorporated.

7.7 Discussion

Driven by the aim to develop better training in coping skills, this paper presented a neurologically inspired computational model of emotion regulation strategies, with an emphasis on learning of strategies over time. The model has some aspects in common with existing models for emotion regulation based on Gross' strategies (Bosse et al., 2010) and fear extinction learning during dreaming (Treur, 2011). The behavior of the model was illustrated by simulation experiments and by mathematical analysis. In addition, a preliminary validation pointed out that the model is able to approximate empirical data obtained from an experiment. Although this is a promising finding, the authors want to stress that this is not an exhaustive investigation that the model is valid. To establish this, more extensive experiments are needed. In addition, we do not claim that there are no other (perhaps simpler) models that are able to replicate the empirical data as closely as the current model. The presented

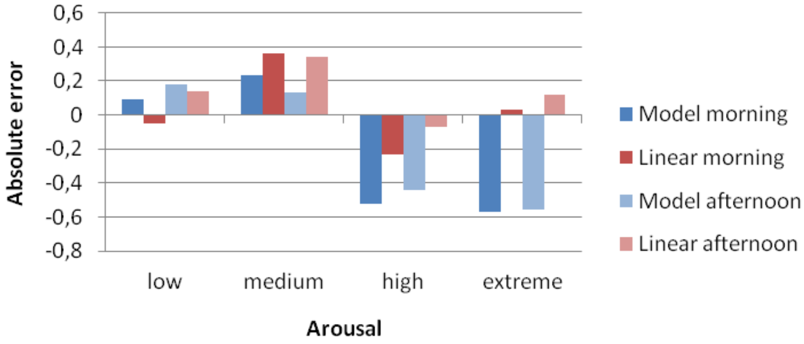

Figure 7.9: Errors in approximated rating for the control group

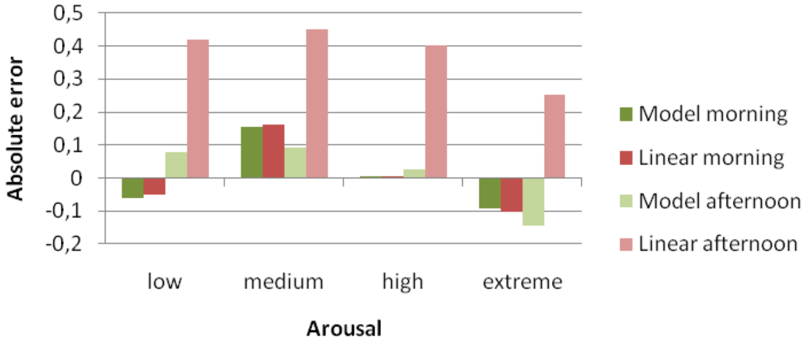

Figure 7.10: Errors in approximated rating for the training group

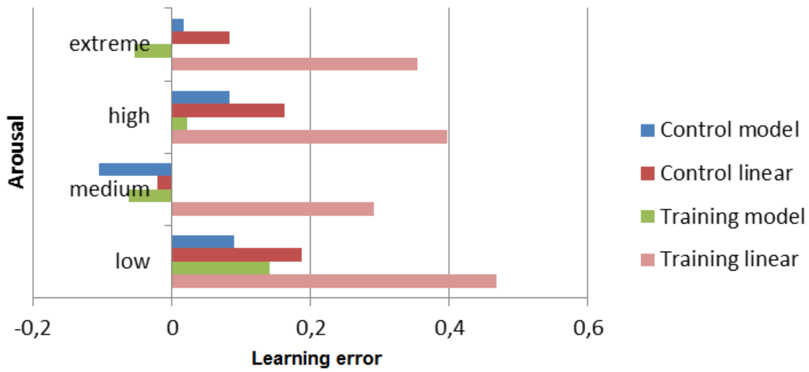

Figure 7.11: Errors in approximating learning effects of emotion regulation

model contains a number of features that have not been fully exploited in the preliminary validation in Section 7.6. However, the choice to develop the model the way we did was not only triggered by the goal to reproduce empirical data (and to build an application), but also by the desire to have a biologically plausible model. In our opinion, this is the main contribution of the model: on the one hand it is based on extensive neurological literature on emotion generation, regulation, and learning, but on the other hand it is sufficiently concrete to be used for real applications. Hence, the expressiveness of the model can only be fully exploited by applying it to a wide variety of case studies. To this end, for future work it is planned to performed more experiments (e.g., making use of different emotion regulation strategies, participants, and time scales), thereby producing a richer set of test data.

Once the model has been extensively tested, a long term goal is to incorporate it in an actual virtual reality-based training environment for professionals in various domains. By having a trainee interact with this environment, measuring her actions and physiological states, and feeding these into the model, the system will be able to reason about the extent to which she has acquired good coping skills, in order to provide appropriate feedback.

Acknowledgments. This research was supported by funding from the National Initiative Brain and Cognition, coordinated by the Netherlands Organisation for Scientific Research (NWO), under grant agreement No. 056-25-013.

References

Becker-Asano, C., & Wachsmuth, I. (2008). Affect simulation with primary and secondary emotions. In *Proceedings of the 8th International Conference on Intelligent Virtual Agents*, (pp. 15–28). Springer.

Berking, M., Meier, C., & Wupperman, P. (2010). Enhancing emotion-regulation skills in police officers: Results of a pilot controlled study. *Behavior Therapy*, *41*(3), 329–339.

Bi, G.-q., & Poo, M.-m. (2001). Synaptic modification by correlated activity: Hebb's postulate revisited. *Annual review of neuroscience*, *24*(1), 139–166.

Bosse, T., Gerritsen, C., De Man, J., & Treur, J. (2012). Measuring stress-reducing effects of virtual training based on subjective response. In *Neural Information Processing*, (pp. 322–330). Springer.

Bosse, T., Pontier, M., & Treur, J. (2010). A computational model based on gross' emotion regulation theory. *Cognitive systems research*, *11*(3), 211–230.

Damasio, A. (1994). *Descartes' Error: Emotion, Reason and the Human Brain*. papermac.

Dias, J., & Paiva, A. (2005). Feeling and reasoning: A computational model for emotional characters. In *Progress in artificial intelligence*, (pp. 127–140). Springer.

El-Nasr, M. S., Yen, J., & Ioerger, T. R. (2000). Flame fuzzy logic adaptive model of emotions. *Autonomous Agents and Multi-agent systems*, *3*(3), 219–257.

Frijda, N. (1986). *The emotions*. Cambridge University Press.

Gebhard, P. (2005). Alma: a layered model of affect. In *Proceedings of the fourth international joint conference on Autonomous agents and multiagent systems*, (pp. 29–36). ACM.

Gerstner, W., & Kistler, W. M. (2002). Mathematical formulations of hebbian learning. *Biological cybernetics*, *87*(5-6), 404–415.

Gross, J. J. (1998). The emerging field of emotion regulation: an integrative review. *Review of general psychology*, *2*(3), 271.

Gross, J. J. (2001). Emotion regulation in adulthood: Timing is everything. *Current directions in psychological science*, *10*(6), 214–219.

Gujar, N., McDonald, S. A., Nishida, M., & Walker, M. P. (2011). A role for rem sleep in recalibrating the sensitivity of the human brain to specific emotions. *Cerebral Cortex*, *21*(1), 115–123.

Hebb, O. (1949). *The organisation of behavior*. wiley.

Kim, M. J., Loucks, R. A., Palmer, A. L., Brown, A. C., Solomon, K. M., Marchante, A. N., & Whalen, P. J. (2011). The structural and functional connectivity of the amygdala: from normal emotion to pathological anxiety. *Behavioural brain research*, *223*(2), 403–410.

Klein, G. A. (1993). A recognition-primed decision (rpd) model of rapid decision making. In G. Klein, J. Orasanu, R. Calderwood, & C. Zsambok (Eds.) *Decision-making in action: Models and methods*, (pp. 138–147). Ablex Publishing Corporation.

Levin, R., & Nielsen, T. A. (2007). Disturbed dreaming, posttraumatic stress disorder, and affect distress: a review and neurocognitive model. *Psychological bulletin*, *133*(3), 482–528.

Marsella, S. C., & Gratch, J. (2009). Ema: A process model of appraisal dynamics. *Cognitive Systems Research*, *10*(1), 70–90.

Oatley, K., & Johnson-Laird, P. N. (1987). Towards a cognitive theory of emotions. *Cognition and emotion*, *1*(1), 29–50.

Ochsner, K. N., & Gross, J. J. (2005). The cognitive control of emotion. *Trends in cognitive sciences*, *9*(5), 242–249.

Phelps, E. A., Delgado, M. R., Nearing, K. I., & LeDoux, J. E. (2004). Extinction learning in humans: role of the amygdala and vmpfc. *Neuron*, *43*(6), 897–905.

Reisenzein, R. (2009). Emotions as metarepresentational states of mind: Naturalizing the belief–desire theory of emotion. *Cognitive Systems Research*, *10*(1), 6–20.

Soleimani, A., & Kobti, Z. (2012). A mood driven computational model for gross emotion regulation process paradigm. In *Proceedings of the World Congress on Engineering and Computer Science*, vol. 1.

Sotres-Bayon, F., Bush, D. E., & LeDoux, J. E. (2004). Emotional perseveration: an update on prefrontal-amygdala interactions in fear extinction. *Learning & Memory*, *11*(5), 525–535.

Szczygieł, D., Buczny, J., & Bazińska, R. (2012). Emotion regulation and emotional information processing: The moderating effect of emotional awareness. *Personality and Individual Differences*, *52*(3), 433–437.

Treur, J. (2011). Dreaming your fear away: A computational model for fear extinction learning during dreaming. In *Neural Information Processing*, (pp. 197–209). Springer.

Yoo, S.-S., Gujar, N., Hu, P., Jolesz, F. A., & Walker, M. P. (2007). The human emotional brain without sleepa prefrontal amygdala disconnect. *Current Biology*, *17*(20), R877–R878.