Modelling Two Emotion Regulation Strategies as Key Features of Therapeutic Empathy

Juan Martínez-Miranda[(✉)], Adrián Bresó, and Juan Miguel García-Gómez

ITACA Institute, Biomedical Informatics Group, Universitat Politècnica de València,
Camino de Vera s/n, Valencia, Spain
{juama25,adbregua,juanmig}@upv.es

Abstract. Computational models of affective processes have allowed the construction of synthetic characters able to produce empathic behaviours. The use of empathy, as a strategy to enhance engagement and cooperation with human pairs has proved good results in different application domains. Mental care is a particular area where the use of empathic virtual characters would offer several advantages facilitating the self-treatment management. Empathic responses in counselling and psychotherapy differ from "natural" empathy produced in everyday situations. Therapeutic empathy requires an emotional involvement of the therapist with the patient and an emotional detachment for a more objective appraisal of the situation. This paper introduces a model of emotion regulation as the first steps to get therapeutic empathy responses in a virtual agent constructed to support the treatment of major depression. The modelling of two specific strategies of emotion regulation based on Gross theory (cognitive change and response modulation) is described.

Keywords: Affective process modelling · Emotion regulation · Reappraisal · Response modulation · Virtual agents

1 Introduction

The interest in the development of Embodied Conversational Agents (ECAs) as advanced human-computer interaction interfaces has generated a good number of initiatives aimed to construct the underlying mechanisms able to produce more human-like behaviours in those agents. The modelling of the emotional phenomenon, as a basic component of human behaviour, has produced different computational models of emotion that are used to analyse and simulate different aspects of this complex process. Most of these computational architectures of emotion are based on different cognitive and psychological theories influenced by the particular components and phases of the emotional phenomenon that the model tries to represent. Examples of these architectures include FLAME [16] and EMA [27] based on appraisal theories of emotion; WASABI [2] based on dimensional theories of emotion; or the model proposed in [1] which is based on the anatomic approach of emotions (for a deeper discussion refer to [29]).

A common expected benefit from these tools is the construction of more believable ECAs that better engage their human pairs during interaction.

One emotion-related element well studied and commonly used to create better interactive scenarios with ECAs is empathy. Empathic agents have been constructed to achieve a better cooperation and complete longer interaction sessions in different domains, including learning [8], training [34] and clinical applications [3]. Within the clinical domain, a particular area where the use of empathic virtual agents can be particularly beneficial is in the treatment of mental health disorders [24]. Empathy is considered a fundamental aspect in promoting therapeutic change when providing counseling and psychotherapeutic interventions [36] and some studies concluded that empathy accounts for between 7–10 % of the variance in therapy outcome [4].

The modelling of empathic responses in ECAs as virtual assistants to support the treatment of mental health disorders faces some challenges that need to be carefully addressed. For example, in the treatment of major depression, an interactive virtual agent must not display a "pure emotional" empathic behaviours by adopting the same—typically negative—mood of the patient. The disadvantage is that these behaviours can be interpreted as sympathetic expressions of condolence that may imply a sense of unintended agreement with the patient's (negative) views [11]. What is most beneficial from a clinical perspective is not to produce "only" natural empathic reactions as response to the patient's input, but to generate *therapeutic-empathy* responses in the agent.

As mentioned in [38], it is important to distinguish *natural empathy* (experienced by people in everyday situations) from *therapeutic empathy* in order to provide the patients with useful feedback for their particular condition. One of the key differences between natural and therapeutic empathy is the "addition of the cognitive perspective-taking component to the emotional one; the cognitive component helps the therapist to conceptualize the client's distress in cognitive terms" ([38], pp. 594). In other words, a therapist should "assume both the role of an emotional involvement in an interview with a patient and an emotional detachment that allows for a more objective appraisal" ([11], pp. 102); a wrong empathic attitude is generated when the therapist does not to some degree maintain an emotional distance from the patient [39].

In this paper, we describe the modelling of this *perspective-taking component* aimed to produce in a virtual agent the required emotional detachment or emotional distance at specific stages of the interaction with patients with major depression. The theoretical basis of the proposed model lies in J. J. Gross' process model of emotion regulation [18]. In particular, we are modelling two strategies of emotion regulation: *(i) cognitive change* and *(ii) response modulation*. The **cognitive change strategy** is triggered when the patient is reporting a *bad situation* (e.g. low mood level) which in a first step would also produce (empathically) a negative emotion in the virtual agent. Once triggered, the cognitive change strategy seeks for additional information that can change (positively) the significance of the detected situation (e.g. finding a positive tendency in the mood level regarding the reported values in past days) allowing *a more objective*

appraisal. Complementarily, the **response modulation strategy** is used to regulate those negative emotion-expressive behaviours in the agent produced when the cognitive change strategy has not succeed (i.e. there is no information that changes the -*negative*- situation meaning). The suppression of negative expressive behaviour helps the virtual agent in not to convey a sense of condolence that would be counterproductive due to the patient's condition. The implementation of the model have been developed as an extension of the FAtiMA (appraisal-based) computational architecture of emotions [14].

The rest of the paper is organised as follows: in Sect. 2 we put in context our research by introducing the Help4Mood project, which aims to remotely support the treatment of major depression. Then, in Sect. 3 we present in more detail the relevant parts of Gross' emotion regulation theory used in our model, complemented by some related work in the ECAs community. Section 4 presents in detail our model of emotion regulation as the core component able to produce better therapeutic empathy reactions in a virtual agent. Section 5 describes the first steps towards the evaluation of the model and finally, Sect. 6 presents the conclusions and further directions of the presented work.

2 A Virtual Agent to Support the Remote Treatment of Major Depression

The model reported here is part of the work developed in an EC-FP7 research project called Help4Mood (www.help4mood.info). The main aim of the project is the development of an interactive system designed to support the treatment of people who are recovering from major depressive disorder in the community, with the focus being on promoting adherence to the therapy through engaging the patients. The complete Help4Mood system is composed of three main components: a virtual agent (VA), which acts as the main interface with the user; a Decision Support System (DSS) which manages, analyses and summarises the user's daily sessions; and the Personal Monitoring System (PMS) in the form of sensor devices used to collect objective measures from the user, including physical activity and sleep patterns.

The virtual agent has been designed to facilitate the collection of relevant data including subjective measures (by applying standardised questionnaires and guided interviews), and neuropsychomotor measures (by offering tasks for speech input and/or selected games) which are designed to complement the objective measures obtained from the PMS. The collection of the patient data is carried out through daily sessions between the patient and the system; the content of each session varies, with some tasks being carried out every day (such as the Daily Mood Check consisting of a single item measuring overall mood plus four items of the CES-D VAS-VA questionnaire [30]), and some others executed weekly (e.g. the standardised PHQ-9 questionnaire [26], for example). Moreover, the virtual agent also helps users to identify negative thinking (a key characteristic of depressive disorder) and challenge it by adapting a protocol
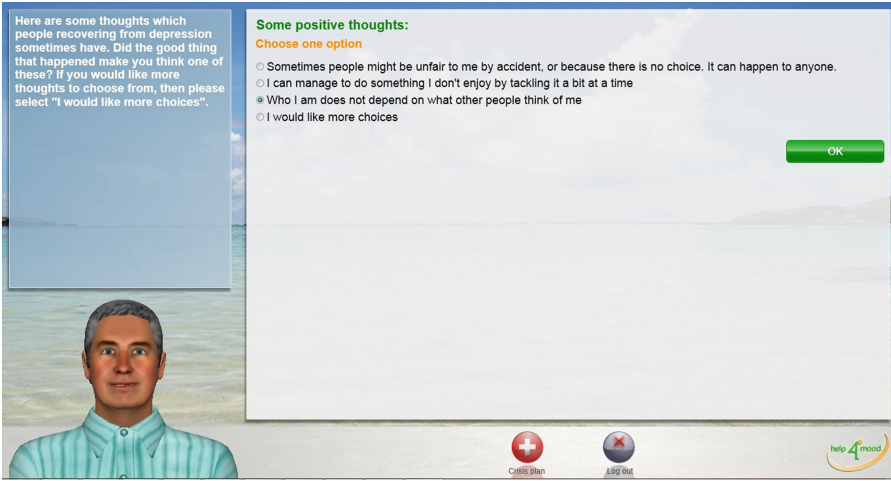
Fig. 1. The Help4Mood GUI

in concordance with the principles of Cognitive Behaviour Therapy (CBT), the main non-pharmacological treatment method for major depressive disorder.

The VA agent is composed by different but inter-related components which process and generate different aspects in the agent's behaviour. The components include the cognitive-emotional module which receives the events inferred in the DSS (using the PMS and user inputs) during a session with the patient. These events are used by the cognitive-emotional module to produce the specific task and the corresponding—if any—emotion to be disclosed during the interaction with the patient. A second component is the Dialogue Manager System (DM) which transforms the task received from the cognitive-emotion module as dialogue acts. The dialogue acts are passed to The Natural Language Generator (NLG), which produces the content of the agent's verbal response in an appropriate style. The generated text string is sent to the Text-To-Speech (TTS) engine which realises the audio with the appropriate tone of voice. The audio and time-aligned phonemes used for lip-syncing are passed to the graphical representation of the virtual agent which takes the form of a talking head immersed in the GUI of the application (see Fig. 1). The current active emotion is also passed to the GUI for the rendering of the corresponding non-verbal communication (i.e. head movements and a set of facial expressions to convey the triggered emotion).

The cognitive-emotional module, which is the focus of this paper, has been developed as a Java-based application which makes use of the FAtiMA architecture [14]. For the Help4Mood scenario, we authored all the goals, actions and action tendencies to cope and react towards the events produced during the interaction with the patient. All these events are directly related with the user's responses and are the basis to produce the next more adequate action and emotion in the VA. As all the actions in the VA are directed to provide the different

standardised questionnaires or CBT-based activities defined by the clinicians, the range of user's responses is delimited to provide the input and follow up of the offered activities. When new questionnaires and/or exercises are necessary to be added -which in turn extend the content of the daily sessions- new goals, actions and action tendencies are authored in the VA to correctly cope with the user's inputs to the new events.

With the aim to promote the usability and acceptability of the Help4Mood solution in the potential users (i.e. patients, clinicians and caregivers), the development of the integrated system has followed the user-centered design methodology [33]. The functionalities of each component are cyclically developed following the feedback and suggestions of the involved users. In the initial usability study of the integrated system[1] (PMS + DSS + VA), we followed the recommendations from the clinical experts of the project. One of them was related with the emotional behaviour in the VA: **it must not convey any negative emotion** during the interaction with the patient. The main reason behind this requirement was to avoid any interpretation of negative emotions as expressions of condolence in the side of the patient that would be clinically counter-productive. Thus, three positive emotions from the OCC model [31] (implemented as the appraisal and affect derivation mechanism in FAtiMA) were used during the interaction:

1. **Joy**: activated when an event is appraised as *desirable* for the VA (e.g. when the user accepts the activities offered by the VA during a session).
2. **Happy-For**: elicited when an event is appraised as *desirable for the patient* (e.g. good self-reported moods, thoughts or scores in the proposed activities).
3. **Admiration**: activated when an event is appraised as a *desirable consequence of a patient's action* (e.g. the correct completion of the proposed activities during the session or the completion of the whole session).

When something *goes wrong* in the clinical condition of the patient (inferred in the DSS from the PMS data or from patient's self-reports), a neutral attitude (i.e. no emotion) is adopted by the VA. The three positive emotions, when elicited, are conveyed to the patient through the combination of open and close mouth smiles plus some head movements such as nodding (identified as a key element to reinforce the sense of attention and understanding during clinician-patient communication [35]). As stated in [15], most of the positive (*enjoyable*) emotions share the smiling expression and it is not straightforward to differentiate them just through the face but other signals, such as the voice, are required. The strategy followed in our scenario was to use the intensity of the triggered emotion to display a mouth open (greater intensity) or mouth closed (lesser intensity) smile. Moreover, the positive emotions are also used by the Dialogue Manager to select and add specific utterances to the verbal feedback provided by the VA (e.g. "*That's great!*" or "*Well done!*").

---

[1] Five people, two males and three females, who recovered from depression were recruited to use the system during a week in daily sessions with a length between 5 to 8 min.

An interview was applied to all the patients at the end of the system's usage to collect their impressions in the utilisation of the hardware (sensors) and software of Help4Mood. The feedback obtained from this initial acceptability study regarding the virtual agent was divided. Two female participants, P2 and P3, liked the virtual agent and found it "cute". P2 also noticed that the virtual agent reacted to her responses on the Daily Mood Check. While P2 thought that the virtual agent's voice was nice and calm, P3 characterised her as sounding upbeat. In contrast, P5, the third female participant, found the voice depressing; she did not pay any attention to the virtual agent and would have preferred the ability to talk to a human via teleconferencing. In terms of the emotional behaviour only 1 of the 5 users noted the emotional reactions in the VA. This result was not much unexpected since the emotional reactions in the VA were displayed only when the participants responded with *good mood*-related values in the Daily Mood Check questionnaire. In all other cases the VA simply asks the next question adopting a neutral stance.

The noted lack of responsiveness in the VA's behaviour is likely to have more than one cause, including the very short interaction time, not much variability in the response dialogues, a not enough adequate emotional intonation in the VA's speech and the aforementioned neutral attitude adopted by the VA in front of negative situations. As more content (i.e. new activities) is added to the daily sessions, longer interactions between the patient and the VA are produced offering a good opportunity to produce more variability in the different aspects of the VA's behaviour.

In terms of the cognitive-emotional component, the production of richer emotional responses became necessary to better engage the users. The inclusion of negative emotions which produce affective reactions to adverse situations would contribute to perceive a more empathic agent. The challenge here is to generate an optimal intensity in these negative emotions that allows an adequate response (in terms of action-selection and feedback provided) during the interaction with the user. To face this challenge, we have incorporated a model of emotion regulation aimed to modulate the negative emotions elicited in the VA. The emotional regulation is achieved following two strategies: changing the perspective of the current situation (producing an emotional detachment), and suppressing the expressive (negative emotion-based) behaviour to convey the appropriate reactions.

3   Emotion Regulation

3.1   Theoretical Foundations

The study and understanding of the emotion regulation process has attracted the interest of an important number of researchers in the last three decades [10,20]. Although some works consider the emotion regulation process as part of the emotion generation process [9,17], some others show the neural differences between them [12] and the benefits of studying emotion and emotion regulation separately [21,25]. In line with the second view, J. J. Gross [18,21] proposed

a theoretical model of emotion regulation which refers to *the heterogeneous set of processes by which emotions are themselves regulated*. In detail, the process model of emotion regulation covers the conscious and unconscious strategies used to increase, maintain, or decrease one or more components of an emotional response. The main characteristic of this model is the identification and definition of five families of emotion regulation processes: *situation selection*, *situation modification*, *attentional deployment*, *cognitive change* and *response modulation*.

Situation selection is described as when an individual takes the necessary actions to be in a situation the individual expects will raise a certain desirable emotion. *Situation modification* refers to the efforts employed by the individual to directly modify the actual situation to alter its emotional impact. The third family, *attentional deployment*, refers to how individuals direct their attention within the current situation in order to influence their emotions. *Cognitive change* is described as when the individual changes how the actual situation is appraised to alter its emotional significance, either by changing how the individual thinks about the situation or the capacity to manage it. Finally, the *response modulation* family refers when the individual influences the physiological, experiential, or behavioural responses to the situation.

Each family of emotion regulation processes occurs at different points in the emotion generation process and there are substantial differences between them (see details in [21]). An important aspect to consider is that the first four emotion regulation families occur before any appraisal produces the full emotional response (antecedent-focused), while the last family (response modulation) occurs after response tendencies have been initiated (response-focused). Two particular strategies of emotion regulation have been studied in [18]: one is *reappraisal* as a type of cognitive change (antecedent-focused) and the other is *suppression* as a type of response modulation (response-focused). According to the authors, reappraisal occurs early in the emotion generation process and it involves cognitively neutralizing a potentially emotion-eliciting situation. In consequence, reappraisal should decrease experiential, behavioural, and physiological responses. On the other hand, suppression occurs later in the emotion generation process and requires an active inhibition of the emotion-expressive behaviour that is generated when the emotion is triggered.

3.2   Computational Models of Emotion Regulation

The Gross process model of emotion regulation has inspired the development of some computational models of emotion regulation. The group of Bosse and colleagues have formally modelled the four antecedent-focused emotion regulation strategies and incorporate it in synthetic characters as participants in a virtual storytelling [6]. In a subsequent work, Bosse and colleagues constructed virtual agents not only with the capacity of regulate their emotions, but also with the ability of *reasoning* about the emotion regulation processes of other agents [5]. This model has been called CoMERG (the Cognitive Model for Emotion Regulation based on Gross) and it formalizes Gross model through a set

of difference equations and rules to simulate the dynamics of Gross' emotion-regulation strategies [7].

CoMERG identifies a set of variables and their dependencies to represent both quantitative aspects (such as levels of emotional response) and qualitative aspects (such as decisions to regulate one's emotion) of the model. These variables include e.g. the level of -the actual- emotion, the optimal -desired- level of emotion, the personal tendency to adjust the emotional value, or the costs of adjusting the emotional value, among others which are used to simulate and evaluate the results in the use of the four antecedent-focus strategies of emotion regulation. The modelling and simulation of the different emotion regulation strategies is the main aim of CoMERG, but the underlying appraisal and affect derivation mechanisms required to generate specific emotions according to the observed world-state are not explicitly addressed. In a more recent work [23] the integration of CoMERG with other two computational models of emotions EMA [27] and I-PEFIC^{ADM} [22] is proposed to cover the complete process of emotion generation, regulation and action responses in virtual agents.

Similarly, the work presented in [37] proposes an extension of CoMERG by adding an emotion-dependent regulation process based on the mood and personality of individuals. Moreover, the occurrence of new (positive and negative) events during the simulation time was included to analyse the influence of these events on the emotion regulation process. However, as an extension of CoMERG, this approach does not have an appraisal and affect derivation mechanism for monitoring events in the world nor have been reported its integration in virtual characters.

It is important to mention that FAtiMA also applies its own strategy (which is based on [28]) for changing world interpretation and lowering strong negative emotions. This mechanism is part of the FAtiMA deliberative layer which implements two types of coping to deal with changes in the environment. The problem-focused coping acts on the agent's world to deal with the situation and consists of a set of actions to be executed to achieve the desired state of the world. The emotional focused coping is used to change the agent's interpretation of circumstances. When a specific plan or action fails in the intention to achieve or maintain a desired goal, a mental disengagement is applied. Mental disengagement works by reducing the importance of the goal, which in turn reduces the intensity of the negative emotions triggered when a goal fails [14].

For the Help4Mood scenario, what is still needed is the mechanism to reinterpret (i.e. reappraise) a situation that is detected as adverse to the patient's condition and that could lead to the triggering of a negative emotion. While the current emotion-focused coping of FAtiMA is concentrated in the achievement/maintenance, or not, of the agent's internal goals and the reduction of the intensity of the negative OCC prospect based emotions (i.e. *disappointment* and *fears_confirmed*), we need an emotional regulation module that down-regulates the intensity of the negative affective state produced by a situation derived from those negative events in the patient's status. Thus, the verbal and non-verbal feedback provided to the patient based on the VA's affective state would

contribute to a better therapeutic empathy communication during the session. The design and implementation of this module is detailed in the following section.

4   Adapting a Model of Two Emotion Regulation Strategies

The initial version of the cognitive-emotional module in the Help4Mood's virtual agent has been extended by allowing the elicitation of two OCC-based negative emotions. *Pity* is activated as a result of appraising some events as *not desirable for the patient* (e.g. when reporting a low mood, negative thoughts or decreased physical activity). *Distress* is triggered when an event is appraised as *not desirable for the agent itself* (e.g. when a not daily use of the system is detected). The challenge is how to communicate these emotions not as a sense of condolence due to the adverse events, but as a sense of understanding and provide useful feedback that motivates the patient towards the daily use of Help4Mood. As with the positive emotions, the negative emotions are reflected through a particular facial expressions and some dialogues constructed in the Dialogue Manager. The main objective is to produce the optimal intensity in the negative emotion to display an adequate facial expression, and at the same time, take the necessary actions to cope with the situation.

The proposed solution is to implement a mechanism of emotion regulation that can be used to modulate the negative emotions and produce an emotional detachment from the situation which helps to provide useful responses during the interaction. Based on Gross theory of emotion regulation, we have implemented two of the strategies defined in Gross's model of emotion regulation: cognitive change (through the reappraisal of events) and response modulation (through suppression of the emotion-expressive behaviour). At the moment, we are not considering the inclusion of the other three antecedent-focused emotion regulation strategies.

The main reason behind this decision is the particular context of the VA's environment in the Help4Mood scenario: the main events received and appraised by the VA are all closely related with the actual detected or reported condition of the patient. During the interaction cycle, there are no other alternative situations to select, i.e. there are no other possible values in the patient's condition (situation selection) that the VA can concentrate on. Also, the VA cannot modify by itself the detected or reported patient condition (situation modification) and it is desirable that the VA should appear focused on what the patient is actually reporting (attentional deployment). Nevertheless, it is still possible to positively *reappraise* the current situation by analysing how the patient's condition has evolved during past sessions. If after the reappraisal process the current situation cannot be assessed in positive terms, some *suppression* in the intensity of the activated negative emotion is still possible to modulate the displayed facial expression and/or head movements related with the current negative affective state of the VA.
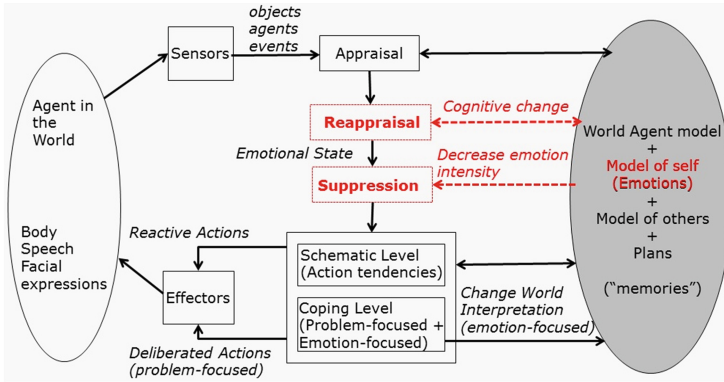
Fig. 2. The FAtiMA architecture [14] with an added emotion regulation component

We have incorporated an initial model of these two emotion regulation strategies as an extension of the FAtiMA architecture. A key advantage of FAtiMA is its modular implementation which is composed of a core functionality plus a set of components that add or remove particular functionalities (in terms of appraisal or behaviour) making it more flexible and easier to extend [13]. Thus, the proposed model of emotion regulation has been added as an extended component of the FAtiMA core functionality as presented in Fig. 2 (the new component is displayed using dotted red lines).

4.1   Modelling Cognitive Change - Reappraisal

Based on Gross theory of emotion regulation, we have implemented a mechanism to reappraise those events susceptible to triggering a negative emotion in the VA. Following the concepts of Gross theory, we represent a situation composed by the event or events produced in the VA's environment. The actual situation meaning can be changed using a pre-defined set of situation meanings which in turn are formed by the different events that are used during the reappraisal process. The reappraisal process is triggered only when the target (negative) emotion exceeds a configured threshold which represents the maximum intensity allowed in the target emotion. The reappraisal process can produce a different -positive-emotion or the same negative emotion with a decreased -*down-regulated*- intensity. In the case where the resultant emotion is still negative with an intensity greater than the desired maximum threshold, the suppression emotion strategy is applied (see next section). The diagram in Fig. 3 graphically represents the different concepts and the flow of the cognitive change process.

According to Gross's theory, the cognitive change is an antecedent-focused strategy of emotion regulation, which means that it occurs before appraisals give rise to full-blown emotional response tendencies [21]. Thus, our model of cognitive change is activated when a new event is received from the environment. A prospective appraisal is executed to assess if the event derives from
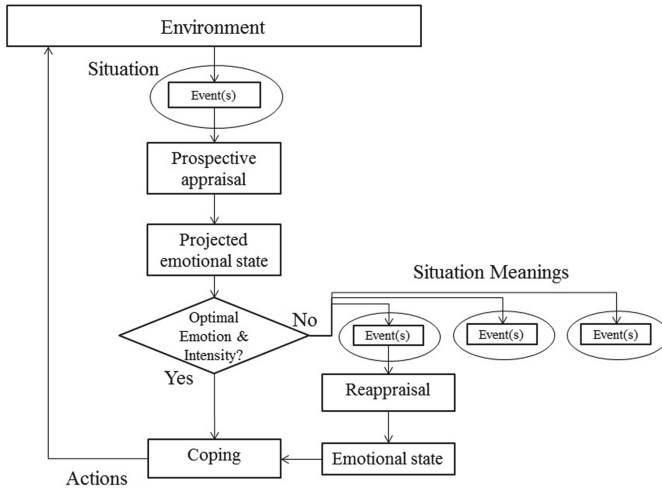
Fig. 3. Cognitive change model diagram

a desirable or undesirable (in terms of the agent's goals) situation related to the patient's condition. The result of this prospective appraisal is the projection of the potential emotional state produced by the event. In other words, our model "simulates" the appraisal and affect derivation processes to analyse the emotional consequences of the current situation, but without producing the full-blown emotional responses.

If the projected emotional state involves the activation of a positive emotion— no emotion regulation is required—then the same event is used to execute the *real* appraisal and generates the corresponding responses in the VA. On the other hand, if the projected emotional state includes the activation of a negative emotion with intensity greater than the pre-defined maximum threshold, the corresponding pre-defined alternative event(s) is selected for reappraisal which would construct a more positive meaning of the original situation. If the emotional state produced by the reappraisal is better (i.e. produces a positive emotion or the same negative emotion with a reduced intensity) than the simulated situation, all the (reactive and deliberative) responses are executed continuing with the next interaction cycle.

To exemplify this process consider the following: during the current Help4Mood session one of the activities to perform is the assessment of the patient's depression level in the last 7 days through the PHQ-9 questionnaire. If the score obtained indicates that the depression level is quite high, the VA can appraise this result as highly undesirable for the patient's condition, generating a strong *pity* emotion. Using the cognitive change strategy of emotion regulation, the VA can change the meaning of this situation using an alternative view. In the example, the VA can consult the results obtained in the PHQ-9 questionnaire during previous sessions (stored in the model of the patient maintained in the DSS module) and check

whether the current result shows a positive tendency in the patient's condition taking into account the previous results. If this positive tendency is found, the original event would be reappraised as "not much undesirable" to the patient (thought the current PHQ-9 score is still not the optimal). This reappraisal can change the emotional state or the emotion's intensity in the VA which is reflected in the feedback provided to the patient, something like "Ok, it seems that your current condition is not very good, but in general terms you are making good progresses in the last days". This is different from the feedback that the VA would provide if the response is based only on the negative meaning of the current situation (e.g. "Ok, it seems that you have had some difficult days, but please continue with the treatment"). In both cases, the verbal feedback is accompanied by the appropriate facial expression according to the activated emotion and its intensity.

Similar events to change the meaning of the current situation can be predefined to cope with the result of other session activities (e.g. the Daily Mood Check questionnaire, the negative thoughts challenge, or the behaviour activation modules). All the targeted emotions to be regulated, the maximum intensity threshold and the different situation meanings with the events used during the reappraisal can be authored in an XML file in a similar fashion as the emotional thresholds and decay rates; the emotional reaction rules; the set of action tendencies; and the goals and actions that form the whole VA's behaviour are currently authored in FAtiMA. A simple example of this XML-based file is as follows:

```
<EmotionRegulation>
   <!--Targeted emotions for regulation-->
   <EmotionalDesiredIntensities>
     <EmotionalDesiredIntensity emotion="Distress" desiredIntensity="3"
                                suppressionRate="3">
     <EmotionalDesiredIntensity emotion="Pity" desiredIntensity="3"
                                suppressionRate="4">
     ...
   </EmotionalDesiredIntensities>

   <!--Situation meanings used for reappraisal-->
   <SituationMeanings>
     <Situation name="High_Depression_Score">
       <ElicitingEmotion type="Distress" minIntensity="6">
         <CauseEvent subject="User" action="High_Score_PHQ-9">
       </ElicitingEmotion>
       </EventForReappraisal subject="[SELF]" action="getPreviousDepressionScore"
                             target="User" parameters="2">
     <Situation>

     <Situation name="Low_Mood">
       <ElicitingEmotion type="Distress" minIntensity="6">
         <CauseEvent subject="User" action="Low_Score_DailyMoodCheck">
       </ElicitingEmotion>
       </EventForReappraisal subject="[SELF]" action="getPreviousMood" target="User"
                             parameters="5">
     <Situation>
   </SituationMeanings>
</EmotionRegulation>
```

The content of the file is divided into two main parts: the first part under the <EmotionalDesiredIntensities> tag defines the targeted emotions that will be regulated (in our case, we concentrate only on negative emotions). It contains the

values for the maximum desired intensity of the emotion and the suppression rate value used in the response modulation strategy (see next section). The second part of the content under the `<SituationMeanings>` tag defines the set of situations that are candidates to be reappraised with a more *positive* perspective. Each situation contains the event (`<CauseEvent>`) that would elicit the *negative* emotion (`<ElicitingEmotion>`) and the definition of the event used for the reappraisal (`<EventForReappraisal>`). The event that is used during the reappraisal process is composed by four fields following the same definition of an event used in FAtiMA [14]:

– The *subject* who performs the action
– The *action* to perform
– The *target* of the action
– A list of parameters that specify additional information about the action

The mechanism to select the specific situation in the emotion regulation component is based on the activation of action tendencies process provided in FAtiMA. When executing the prospective appraisal using the event defined in the `<CauseEvent>` tag, if the projected emotional state activates the emotion type with an intensity equal or greater than the *minIntensity* value, then the event defined in `<EventForReappraisal>` is selected for the reappraisal. This event contains the action that will be executed to get an alternative meaning of the current situation (e.g. in the *Low_Mood* situation of the XML example, the virtual agent gets the scores of the mood reported by the patient in the past 5 days to detect if the mean of the previous values is high or if there is a positive tendency in the mood of the patient).

The result of the reappraisal does not necessarily change a negative situation. Continuing with our example, the result of the previous depression scores could indicate a negative tendency in the evolution of the patient. In these cases, the resultant emotional state could even increase the intensity of the negative emotion. As a requirement of the VA in the Help4Mood scenario is not to convey strong negative emotional responses during the interaction, the response modulation strategy is activated in these cases.

4.2   Modelling Response Modulation - Suppression

Response modulation is an emotional regulation strategy that occurs late in the emotion-generative process, once the response tendencies have been initiated [21]. According to [19], a common form of response modulation involves regulating emotion-expressive behaviour. In this sense, suppression as a form of response modulation can be used to model the regulation of the activated ongoing expressive behaviour in the VA during the interaction with the patient. The activation of negative emotions in the VA could be useful in selecting the appropriate action to cope with specific situations. An example that can occur during the daily sessions is when the VA detects cues that indicate thoughts of self-harming (i.e. a high score in question number 9 of the PHQ-9 questionnaire).

In this case, an action tendency triggered by the activated -negative- emotional state is the execution of the Help4Mood crisis plan: display contact information for crisis services and trusted family/friends plus discontinuing the use of Help4Mood. The expressive behaviour (i.e. the displayed facial expression and the verbalised utterances) during the execution of the crisis plan should avoid an unnecessary sense of alarmism, but promote calming and understanding of the situation.

The way to regulate the expressive behaviour in the VA produced by the on-going emotional state is through decrementing the intensity of the emotion. As the intensity is used to generate the corresponding facial expression (stronger intensity means more marked expressivity in the face of the VA), what we need is a mechanism to down-regulate the intensity of the activated negative emotion. In FAtiMA, the intensity of an emotion is a variable which depends on the elapsed time and it is influenced by a pre-defined decay rate parameter. The decay function for each emotion is implemented in FAtiMA using the following formula, as proposed in [32]:

$$Intensity(em, t) = Intensity(em, t0) \cdot e^{-bt} \tag{1}$$

Where the *Intensity* of an emotion *(em)* at any time t depends on the intensity of the emotion when it was generated *Intensity(em, t0)* and the value of the decay rate b determines how quickly the intensity of the emotion decays over time. A slight modification of this formula by introducing an additional value called suppression rate sR is used to introduce a high decrement in the emotion intensity aiming to reach the optimal intensity value in the regulated emotion:

$$Intensity(em, t) = Intensity(em, t0) \cdot e - bt \cdot \frac{1}{sR} \quad \text{where } sR >= 1 \tag{2}$$

The two values, the optimal intensity and the suppression rate, are pre-configured for each emotion in the emotion regulation XML-based file presented in the previous section. Different values for these parameters will generate different behaviour in the VA: higher values in the *desiredIntensity* parameter will allow stronger intensities in the negative emotions triggering the corresponding coping behaviour. In contrast, low values for this parameter will force the application of the implemented emotion regulation strategies to achieve the desired intensity in the negative emotions. On the other hand, higher values in the *suppressionRate* parameter will produce a faster decrement of the emotion intensity and the suppression of the emotion expressive behaviour.

5   Evaluation

The evaluation of the proposed model is currently ongoing within the evaluation of the whole Help4Mood system through the running of the final pilot involving patients recruited by the clinicians of the consortium. The inclusion criteria for the participants include those patients with major depressive disorder as a

Table 1. Questions for the assessment of the agent interaction

Number	Question	Anchor 1	Anchor 5
Q1	The virtual agent behaves cold and aloof	Strongly disagree	Strongly agree
Q2	The virtual agent looks emotionally stable	Strongly disagree	Strongly agree
Q3	I am comfortable with the emotional responses of the VA	Strongly disagree	Strongly agree
Q4	The virtual agent is trustworthy	Strongly disagree	Strongly agree
Q5	I would like to continue interacting with the virtual agent	Strongly disagree	Strongly agree
Q6	The virtual agent motivates me to use the Help4Mood system on daily basis	Strongly disagree	Strongly agree

primary diagnosis with a mild to moderate range[2]; aged between 18 and 64 inclusive; and living at home. Excluded participants are those with a recent major adverse life event such as bereavement; a history at any time of other disorders such psychotic depression, bipolar disorder or substance misuse; difficulties in comprehension, communication or dexterity; requirement of personal assistance for activities of daily living; or people whose antidepressant medication had been changed within three months prior to enrolment.

The participants will use the Help4Mood system for 4 weeks followed by an exit interview containing questions to assess the acceptability of the system's components. Regarding the assessment of the VA's acceptability, six (Likert-based scale) questions were designed to collect the feedback from the participants (see Table 1). The first three questions are directly related with the observed emotional behaviour of the VA. These questions will help to assess whether the users perceive an adequate empathic behaviour from the VA and analyse if in the long run it could contribute to a better engagement and long-term use of the whole system as part of the treatment.

A minimum of 15 participants are expected to be enroled in the final pilot which is already ongoing. At the moment, the feedback from the first three patients has been received. All three participants were female aged 38(P1), 42(P2), 58(P3) and used to the use of technology (laptops, smartphones). Two of the three participants (P2 and P3) rated as *disagree* Q1 stating that they did not consider the VA as cold and aloof; P1 rated Q1 as *agree*. Regarding Q2, P2 and P3 rated *agree* while P3 rated *neither agree or disagree*. The responses to Q3 were similar while P2 and P3 rated *agree*, P1 rated it as *disagree*. Similar responses were obtained in Q4, Q5 and Q6, while P2 and P3 consider the VA as trustworthy, reported that they like to continue the interaction with the VA and

---

[2] Measured through standardised questionnaires such as the Beck Depression Inventory II.

with the daily use of the system, P1 responded more negative to these questions (and it was confirmed that she did not logged into the system on daily basis during the two weeks). Moreover, during the exit interview it was found that the low level of acceptance from P1 to the VA was related not only to the VA's behaviour but also to its appearance. This is an important consideration due that a negative view of the VA's appearance would influence also negatively in the perception of its behaviour (and vice-versa). Although the design of the VA's different appearances (male and female characters wearing formal and informal clothes) was performed following suggestions, preferences and recommendations from users and clinicians, it is important to study to what extent how much of the positive/negative ratings on VA's behaviour is related with positive/negative rating on VA's appearance once all the feedback from the pilot's participants is collected. Despite the different ratings about the VA from the first three participants, all of them rated as positive the use of the whole Help4Mood system as a tool to complement the clinicians with the remote follow-up of the treatment.

Although these initial results are interesting, the small number of participants that already provided the feedback is not relevant to get definite conclusions. These initial results will be complemented with the rest of the participants in the final pilot that is currently ongoing. Nevertheless, what is interesting is that with the inclusion of the emotion regulation model, the participants in the second pilot noted better the emotional reactions from the VA than the participants in the first pilot. This suggests that the inclusion of negative (but regulated) emotional reactions in the VA to the reported adverse events in patient's wellbeing contributes to better convey *adequate* empathic reactions.

6   Conclusions and Further Work

The combination of the two—reappraisal and suppression—emotion regulation strategies produce more varied emotional responses in the Help4Mod's VA. In particular, the emotional reactions of the VA in front of *adverse situations* have been improved and facilitates the provision of a more empathic feedback according to the detected events. Initial tests have been performed to analyse the different reactions and feedback produced during the reappraisal of some negative events. These new emotional reactions has facilitated the inclusion of more specific dialogues during the session which in turn would facilitate a better level of acceptability in the users.

Nevertheless, the significant evaluation of the model is expected at the end of the final pilot where the feedback from all the participants will be collected. Similarly to the first pilot, an interview to all the participants will be administered to acquire relevant findings that help with the improvement of the system in a next development phase of the project. It is expected that the feedback obtained from the participants in the final pilot will support to better assess whether the believability and acceptability of the virtual agent has increased.

In terms of further work, the current presented model can be extended in at least one interesting direction: the inclusion of a mechanism to select the specific

emotion regulation strategy depending on the personality modelled in the virtual agent. During the user and system requirements stage of the Help4Mood project, some of the people in the group of potential users identified the importance to get two different styles of interaction in the virtual agent. While a group of users prefer a closer and friendly virtual agent, some others suggest that the virtual agent should adopt a more formal or professional stance.

At the moment, these two different *personalities* have been modelled by authoring different thresholds and decay rates in the modelled emotions. The thresholds for the activation of the positive emotions in the *friendly* virtual agent are smaller than in the *formal* version of the agent. Moreover, the decay rates for these positive emotions in the friendly agent are also smaller than in the formal agent which produce more positive emotional reactions in the first one and a more neutral attitude in the second one. What it is interesting to model in terms of emotion regulation, is the selection of the specific emotion regulation strategy and its frequency of use based on the different personalities. There is an evidence that the emotion regulation strategies habitually adopted by people are related with some individual differences characterised by different personalities [25]. The behaviours produced in the agent through the selection and frequency of the particular emotion regulation strategy would help to clearly differentiate the two personalities and provide the users with their preferred style of interaction.

An additional interesting further work is to investigate whether the regulation of negative emotions is enough to produce useful therapeutic empathy responses. At the moment, and following clinicians recommendations, we have concentrated on the regulation of negative emotions. Depending on the results collected from the final pilot, we would assess if there would be situations where even when the user is reporting a good input to a specific question, the VA should regulate its positive emotional responses reflecting on a more general assessment of current patient's condition.

References

1. Armony, J.L., Servan-Schreiber, D., Cohen, J.D., LeDoux, J.E.: Computational modeling of emotion: explorations through the anatomy and physiology of fear conditioning. Trends Cogn. Sci. **1**, 28–34 (1997)
2. Becker-Asano, C.: WASABI: affect simulation for agents with believable interactivity. Ph.D. thesis for the Faculty of Technology, University of Bielefeld, Bielefeld, Germany (2008)
3. Bickmore, T., Puskar, K., Schlenk, E., Pfeifer, L., Sereika, S.: Maintaining reality: relational agents for antipsychotic medication adherence. Interact. Comput. **22**(4), 276–288 (2010)

4. Bohart, A., Elliott, R., Greenberg, L.S., Watson, J.C.: Empathy. In: Norcross, J.C. (ed.) Psychotherapy Relationships that Work, pp. 89–108. Oxford University Press, New York (2002)

5. Bosse, T., de Lange, F.P.J.: Development of virtual agents with a theory of emotion regulation. In: Proceedings of IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology, WI-IAT'08, vol. 2, pp. 461–468 (2008)

6. Bosse, T., Pontier, M., Siddiqui, G.F., Treur, J.: Incorporating emotion regulation into virtual stories. In: Pelachaud, C., Martin, J.-C., André, E., Chollet, G., Karpouzis, K., Pelé, D. (eds.) IVA 2007. LNCS (LNAI), vol. 4722, pp. 339–347. Springer, Heidelberg (2007)

7. Bosse, T., Pontier, M., Treur, J.: A computational model based on Gross' emotion regulation theory. Cogn. Syst. Res. **11**, 211–230 (2010)

8. Burleson, W., Picard, R.: Affective agents: sustaining motivation to learn through failure and state of stuck. In: Proceedings of workshop of Social and Emotional Intelligence in Learning Environments, Brazil (2004)

9. Campos, J.J., Frankel, C.B., Camras, L.: On the nature of emotion regulation. Child Dev. **75**, 377–394 (2004)

10. Campos, J.J., Campos, R.G., Barrett, K.C.: Emergent themes in the study of emotional development and emotion regulation. Dev. Psychol. **25**, 394–402 (1989)

11. Clark, A.J.: Empathy in Counseling and Psychotherapy: Perspectives and Practices. Lawrence Erlbaum Associates, Mahwah (2007)

12. Davidson, R.J., Fox, A., Kalin, N.H.: Neural bases of emotion regulation in nonhuman primates and humans. In: Gross, J.J. (ed.) Handbook of Emotion Regulation, pp. 47–68. The Guilford Press, New York (2007)

13. Dias, J., Mascarenhas, S., Paiva, A.: FAtiMA modular: towards an agent architecture with a generic appraisal framework. In: Proceedings of the International Workshop on Standards for Emotion Modeling (2011)

14. Dias, J., Paiva, A.C.R.: Feeling and reasoning: a computational model for emotional characters. In: Bento, C., Cardoso, A., Dias, G. (eds.) EPIA 2005. LNCS (LNAI), vol. 3808, pp. 127–140. Springer, Heidelberg (2005)

15. Ekman, P.: Emotions Revealed: Recognizing Faces and Feelings to Improve Communication and Emotional Life. Times Books, New York (2003)

16. El Nasr, M.S., Yen, J., Iorger, T.: FLAME: fuzzy logic adaptive model of emotions. Auton. Agent Multi-Agent Syst. **3**(3), 219–257 (2000)

17. Frijda, N.H.: The Emotions. Cambridge University Press, Cambridge (1986)

18. Gross, J.J.: Emotion regulation in adulthood: timing is everything. Curr. Dir. Psychol. Sci. **10**(6), 214–219 (2001)

19. Gross, J.J., Richards, J.M., John, O.P.: Emotion regulation in everyday life. In: Snyder, D.K., Simpson, J.A., Hughes, J.N. (eds.) Emotion Regulation in Couples and Families: Pathways to Dysfunction and Health, pp. 13–35. American Psychological Association, Washington DC (2006)

20. Gross, J.J. (ed.): Handbook of Emotion Regulation. The Guilford Press, New York (2007)

21. Gross, J.J., Thompson, R.A.: Emotion regulation: conceptual foundations. In: Gross, J.J. (ed.) Handbook of Emotion Regulation, pp. 3–24. The Guilford Press, New York (2007)

22. Hoorn, J.F., Pontier, M., Siddiqui, G.F.: When the user is instrumental to robot goals. First try: agent uses agent. In: Proceedings of IEEE/WIC/ACM Web Intelligence and Intelligent Agent Technology WI-IAT 2008, vol. 2, pp. 296–301 (2008)

23. Hoorn, J.F., Pontier, M., Siddiqui, G.F.: Coppélius' concoction: similarity and complementarity among three affect-related agent models. Cogn. Syst. Res. **15–16**, 33–49 (2012)
24. Hudlicka, E., Lisetti, C., Hodge, D., Paiva, A., Rizzo, A., Wagner, E.: Panel on artificial agents for psychotherapy. In: Proceedings of AAAI Spring Symposium: Emotion, Personality, and Social Behavior'08, pp. 60–64 (2008)
25. John, O.P., Gross, J.J.: Individual differences in emotion regulation. In: Gross, J.J. (ed.) Handbook of Emotion Regulation, pp. 351–372. The Guilford Press, New York (2007)
26. Kroenke, K., Spitzer, L., Williams, J.: The PHQ-9: validity of a brief depression severity measure. J. Gen. Intern. Med. **16**(9), 606–613 (2001)
27. Marsella, S., Gratch, J.: EMA: a model of emotional dynamics. Cogn. Syst. Res. **10**(1), 70–90 (2009)
28. Marsella, S., Gratch, J.: Modeling coping behavior in virtual humans don't worry, be happy. In: Proceedings of the 2nd International Joint Conference on Autonomous Agents and Multiagent Systems - AAMAS'03 (2003)
29. Marsella, S., Gratch, J., Petta, P.: Computational models of emotion. In: Scherer, K.R., Bänziger, T., Roesch, E.B. (eds.) Blueprint for Affective Computing: A Sourcebook, Chap. 1.2, pp. 21–46. Oxford University Press, Oxford (2010)
30. Moullec, G., Maïano, C., Morin, A.J.S., Monthuy-Blanc, J., Rosello, L., Ninot, G.: A very short visual analog form of the Center for Epidemiologic Studies Depression Scale (CES-D) for the idiographic measurement of depression. J. Affect. Disord. **128**(3), 220–234 (2010)
31. Ortony, A., Clore, G., Collins, A.: The Cognitive Structure of Emotions. Cambridge University Press, Cambridge (1988)
32. Picard, R.: Affective Computing. MIT Press, Cambridge (1997)
33. Preece, J., Rogers, Y., Sharp, H.: Interaction Design: Beyond Human-Computer Interaction. Wiley, New York (2002)
34. Prendinger, H., Ishizuka, M.: The empathic companion: a character-based interface that addresses users' affective states. Appl. Artif. Intell. **19**, 267–285 (2005)
35. Rogers, C.: Your body language speaks loudly: nonverbal communication makes patient more comfortable. In: Academy News of The American Academy of Orthopaedic Surgeons. http://www2.aaos.org/acadnews/2002news/b16-7.htm
36. Rogers, C.R.: On Becoming a Person: a therapist's view of psychotherapy. Constable and Company, London (1967)
37. Soleimani, A., Kobti, Z.: A mood driven computational model for gross emotion regulation paradigm. In: Proceedings of the 5th International C* Conference on Computer Science and Software Engineering, pp. 9–17. ACM, New York (2012)
38. Thwaites, R., Bennett-Levy, J.: Conceptualizing empathy in cognitive behaviour therapy: making the implicit explicit. Behav. Cogn. Psychother. **35**, 291–612 (2007)
39. Vanaerschot, G.: The process of empathy: Holding and letting go. In: Liataer, G., Rombauts, J., Van Balen, R. (eds.) Client-Centered and Experiential Psychotherapy in the Nineties, pp. 269–293. Leuven University Press, Leuven (1990)