

Document downloaded from:

<http://hdl.handle.net/10251/81732>

This paper must be cited as:

Bresó Guardado, A.; Martínez-Miranda, J.; Botella Arbona, C.; Baños Rivera, RM.; García Gómez, JM. (2016). Usability and acceptability assessment of an empathic virtual agent to prevent major depression. *Expert Systems*. 33(4):297-312. doi:10.1111/exsy.12151.



The final publication is available at

<http://dx.doi.org/10.1111/exsy.12151>

Copyright Wiley

Additional Information

Assessment of Usability and Acceptability in a User-Adapted System Addressing Depression and Suicide

A. Bresó^{1*}, J. Martínez-Miranda², C. Botella^{3,4}, R. M. Baños^{4,5}, J. M. García-Gómez^{1,6}

¹ Grupo de Informática Biomédica, Instituto ITACA, Universitat Politècnica de València, Spain
adbregua@upv.es, juanmig@upv.es

² Centro de Investigación Científica y de Educación Superior de Ensenada, Unidad de Transferencia Tecnológica (CICESE-UT3), Tepic, México
jmiranda@cicese.mx

³ Universitat Jaume I, Castellón, Spain
botella@uji.es

⁴ Ciber Physiopathology Obesity and Nutrition (CB06/03 Instituto Salud Carlos III Health Institute, Madrid, Spain)

⁵ Universitat de València, Spain
banos@uv.es

⁶ Unidad Mixta de Investigación en TICs aplicadas a la Reingeniería de Procesos Sociosanitarios (eRPSS), Instituto de Investigación Sanitaria del Hospital Universitario y Politécnico La Fe, Valencia, Spain

Keywords: Human-Computer Interaction; user-adapted sessions; emotional virtual agent; usability; acceptability.

*Corresponding author

Abstract

In Human-Computer Interaction, the adaptation of the content and the way of how this content is communicated to the users in interactive sessions is a critical issue to promote the acceptability and usability of any computational system. So we present a user-adapted interactive platform addressing depression and suicide symptoms, in order to evaluate the user acceptance and the system usability.

An empathic Virtual Agent is the main interface with the user and it has been designed to generate the appropriate dialogues and emotions during the interactions according to the detected user's specific needs. This personalization is based on a dynamic user model nurtured with clinical, demographical and behavioral information.

The evaluation was performed with 60 university members. The obtained results were promising, allowing the execution of a further clinical trial. The system's usability was 75.7%, and the user-adapted content and the emotional responses of the VA 70.9%.

1. Introduction

Human-Computer Interaction (HCI) is becoming ever more important in interactive systems for human-centered computing. This cross-discipline (including e.g. engineering, psychology, ergonomics, and design) investigates and tackles all issues related to the design and implementation of better systems for the interaction between humans and computers. Good design of HCI systems enhances the quality of the interaction between people and computers. It is responsible to achieve a good usability level (Dix, et al., 1993), ease of use, acceptable productivity and safety of the system (Preece, et al., 1994). A key issue in the good design and implementation of interactive systems is the assessment of the user conditions in order to adapt the contents and style of interaction to the individual characteristics, skills and behaviors detected in each person (e.g. age, gender, personality, etc.) (Wirtz, et al., 2009) (Darves, et al., 2002) (Rajamanickam, 2011) (Capuano, et al., 2015) (Arning & Ziefle, 2007)

A good interactive system, should not only adjust its contents to the user's characteristics, but also dynamically update these contents at each cycle of the interaction depending on the user responses. These capabilities are known as adaptive HCI. Hence, unlike passive interfaces, an adaptive HCI can adapt its appearance, its content or its structure. **Adaptive HCI** can be considered a complex system, that need to correctly react to the different inputs data, such as visual (e.g. face recognition, gesture recognition, body movement tracking, or eye movement tracking), audio (e.g. speech recognition, voice-based, emotion analysis, or noise detection), traditional (e.g. keyboard, mouse, or touch screen), or physiological (e.g. position, humidity, respiration, pressure, temperature, or galvanic response) inputs. All these different types of inputs need to be processed and correctly interpreted to get a suitable profile of the user's characteristics, capabilities and limitations, in order to "*say the 'right' thing at the 'right' time in the 'right' way*" (Fischer, 2001). So the design of an internal user representation, i.e. a **user model**, is required in personalized interactive systems. Attending to the data gathering method and the design process, two classifications of a user model are considered. Basically, a user model can be dynamic or static, depending on whether it contains a continuous adaptation or not based on the information provided by the user. A second capability is based on asking or learning processes depending on whether it acquires the information through direct questions to the user or observing and interpreting the behavior and actions of the user (Johnson & Taatgen, 2005).

Adaptive systems are particularly required in clinical contexts, specifically when the target users may be people suffering from a mental disorder, such as depression or suicidal tendencies. Both, depression and suicide are two major problems in the global society and are close related. A study shows that 90% of adolescents who die by committing suicide had a mental health problem, usually depression, substance abuse or both (Bole Williams, 2013). The World Health Organization (WHO) estimated that 804.000 suicide deaths occurred worldwide in 2012 (Organization, 2014) becoming a leading cause of death among college and university students in the United States (Schwartz, 2006). Specifically, suicide rates are the second leading cause of death among people between 15–29-year-olds globally. Other relevant data is that for each adult who committed suicide there might be more than 20 others attempting it. Unfortunately, many people experience the first symptoms of depression during their college or university years (Pelkonen, et al., 2005). If these symptoms are not opportunely detected and the individual does not receive an appropriate support, in some cases the person commits suicide. Early diagnosis and treatment of depression can relieve depression symptoms, prevent depression from returning, and help students succeed in college and university (Cuijpers, et al., 2012).

The interactive system described in this paper has been developed attending a user-adapted HCI approach in order to ensure personalized and adaptive interaction with users for the early detection of symptoms related with depression and/or suicidal behaviour. This system is based on an empathic virtual agent that continuously interacts with the user. Depending on the detected condition of the user, the virtual agent provides (1) a set of personalized activities, and (2) a set appropriate facial expressions and dialogues

representing an emotional state in the VA to convey empathic responses while interacting. The system has been evaluated in three academic environments in order to assess its level of acceptability and usability among the users' community. The evaluation has been performed attending three main goals: (1) confirm that the system was deemed appropriate from an ethical point of view; (2) assess if the content and general structure of the sessions included in the system were appropriate and motivated the users; and (3) evaluate the behavior and the emotions management of VA.

The rest of the paper is structured as follows: section 2 presents an overview of existing adaptive HCI systems. Section 3 presents our framework while section 4 describes the adaptation of our framework in the prevention of suicide and depression in the community of University student's scenario. This adaption allows us to perform an evaluation of the user acceptability and system usability. Finally, we discuss about findings and the further work in section 5.

2. Related Work

The identification and prevention of problems related with mental disorders are a clinical research hot topic. There is a lot of literature and many initiatives aimed to prevent cases of suicide, specifically in universities (Corrieri, et al., 2014) (Kirsch, et al., 2014) (Joffe, 2008) (Haas, et al., 2003). In the last years there is an increasing number of universities implementing evidence-based programs designed to identify and prevent distress-related problems and suicide behaviours in young adults. Several universities, such as the *University of Cambridge*, provide special services related to support and guide to people with suicidal intentions. The *Universidad Autónoma de Madrid* published a protocol (Torre Martí, 2013) for suicide prevention and for supporting users in the initial steps, as students experience the first signs or symptoms. The Ohio State University founded a suicidal prevention training for teaching faculty, staff and students including issues about how to (1) recognize warning signals, (2) engage with empathy, (3) ask directly about suicide, (4) communicate hope, and (5) help suicidal individuals to access formal care and treatment.

The continuous and wide use of Information and Communication Technologies (ICT) in the universities has promoted the emergence of technological-based solutions adapted to the problems of these communities. There are a large number of software systems addressed to the prevention of depression and suicide in the university community. Griffiths et al. (Griffiths & Christensen, 2007) present an evaluation of the effectiveness and efficacy of two unguided internet-based self-help systems (commonly used by University students): MoodGYM (<http://moodgym.anu.edu.au>) and BluePages (<http://bluepages.anu.edu.au>). MoodGYM is a web-based system that provides five interactive training modules in order to increase the knowledge about the symptoms of depression, negative automatic thoughts, dysfunctional attitudes, emotions and coping strategies to face stress and interpersonal relationships. BluePages provides evidence-based information about depression including information about symptoms; general and specific sources of help; information about the effectiveness of medical, psychological and alternative treatments; and information about how to prevent depression. Griffiths concluded asserting that both programs are associated with improvements in mental health, knowledge and the promotion of positive attitudes in the users. MEMO is another technological-based solution presented by Whittaker et al. (Whittaker, et al., 2012) as a cognitive behavioral therapy (CBT) based intervention to prevent depression in adolescents via mobile phone texts and video messages. They evaluated the acceptability and utility of the intervention through a randomized controlled trial (RCT) enrolling a total of 1348 volunteered students. In general, the obtained results were positive and the majority of the participants (82.4%) perceived the system as helpful and useful. The system succeeded in improving positive thoughts, promoting the solution of problems, having fun, and dealing with particular issues in the school. More than 1,200 American universities and colleges have teamed up with the Jed Foundation (www.jedfoundation.org) for the development of the *ULifeline* project (www.ulifeline.org), an online resource center for college students that provides a self-assessment

questionnaire and information on suicide prevention and mental health issues such as depression. Jed Foundation provides other similar online resource center called *Half of us* (www.halfofus.com) that aims to help students and their friends by Bueasasasas in order to improve the interaction with the users. Manning and colleagues (Manning, et al., 2012) (Tarashankar, et al., 2012) used a virtual agent to manage the stress of the students during the exams period. High levels of stress have been reported during exams period that can lead to depression or suicide (among other mental health implications for students) (Lee, 2010). Offering adequate support during these periods helps to prevent depression and suicide. In the Manning and colleagues' work, the students verbally interact with a highly personalized and animated virtual character called *eCounselor*. The *eCounselor* takes the roll of a virtual psychologist, which is implemented through an embodied conversational agent (ECA) with the modeling of behavioural traits, a set of facial expressions and gestures that make the interaction more immersive. The *eCounselor* aims to enhance the student coping strategy and converting his stress into work efficiency. Based on a user model, the *eCounselor* tries to maximize personalization and anthropomorphism. Personalization refers to providing tailored exam management advice to accommodate individual differences. Anthropomorphism (Zanbaka, et al., 2006) refers to the extent a virtual agent resembles the appearance or behavioural attributes of a human psychologist. A similar approach is used in the Miami University containing two avatar-based programs developed by *Kognito Interactive* (www.kognito.com): "*At-Risk for College Students*" and "*At-Risk for Faculty & Staff*" (Bryan & Albright, 2013). Both are online role-play training simulations that help to identify emotionally distressed students and faculty members and prevent suicide. The users enter into a virtual environment and engage in a series of interactive exercises including a simulated conversation with a virtual student that exhibits signs of psychological distress, anxiety, depression, and/or suicidal ideation. A sample of 270 students from 10 different American states was included in the evaluation. A significant increment of acquired skills to manage the symptoms was reported after a 3-month duration study.

In general, the reported findings from the online counselling applications have positively evaluated in the promotion of working alliance, helpfulness and impact and most of them have reported client improvement and satisfaction (Richards, 2009). Regarding the use of virtual agents, there are supporting studies suggesting that people usually feel more comfortable when interviewed by media-mediated electronic doctors and are more likely to release their flinched mind during the consultation process in contrast to human doctors (Yoshida, et al., 2002). Nevertheless, in order to optimize the effectiveness of these tools, a high level of personalization is required during the course of a session and through every session. The next section presents the components of a modular framework as the core component of an interactive system aimed to dynamically adapt both, the contents and the style of interaction offered through a virtual agent to prevent symptoms of depression and suicide ideation in students.

3. Framework

The framework evaluated in this work is composed of four main layers. Two of these layers are used to personalize the contents of the session: (1) a **cognitive layer**, which is the module responsible of the processing of user's incoming information, knowledge extraction and updating of the user's model; (2) a **session planner layer**, the module that selects and plan the contents to be offered to the user during a session according to the information inferred in the cognitive layer. Two additional layers are used to personalize the style of the interaction, i.e. the most adequate way to communicate the contents to the user. These two layers are (3) an **emotional layer** responsible to produce an adequate and coherent emotional state in the virtual agent in order to behave empathically towards the user; and a (4) **visualization layer** which is the graphical representation of the virtual agent and used to convey the produced emotions in the emotional layer through a set of facial expressions.

The four layers of the framework have been designed to be independently from each other and the input/output information of each layer is XML-based formatted facilitating a highly configurable system

for different case of studies (e.g. the provision of support for other mental health related problems such as anxiety or phobias). The functionality of each layer has been conceived to produce the output information according to the user's model updated from the user's inputs and based on some clinical pre-defined guidelines assuring a high level of adaptation to the personal characteristics of each user.

3.1. The Cognitive Layer

The core of this layer is a Rule Based System (RBS) configured with multiple "if-then" rules used to codify some clinical-based expert knowledge. The rules allow to infer (1) the condition of the user based on the responses to direct questions, (2) the updating of the user's model through the addition or change in values of clinical-related concepts (which are coded using an internal format based on the SNOMED-CT¹ terminology), (3) the selection of a set of activities and questionnaires based on Cognitive Behaviour Therapy (CBT) that can be offered to the user, and (4) the generation of alerts if the system detects a high risk situation in the user, such as suicide ideation.

The cognitive layer is the component that maintains a continuous updating of the user's model according to the inferences made over the new input information. The data that form the model of the user is presented in Figure 1, including objective (e.g. obtained questionnaires scores, user selections, rejected and completed tasks, user logins, elapsed time, or number of critical situations) and subjective information (assessment of each user response). Depending on the user's inputs at each interaction cycle, the selection of the action(s) (e.g. recommend some activities or ask other questions to gather more information) that the virtual agent will implement in the next interaction cycle is executed. For example, according to the user response to the Patient Health Questionnaire 9 (PHQ-9), a standardized questionnaire to assess the level of depression, the RBS would infer based on the SNOMED-CT terminology, a new condition in the user (*mild depression {310495003}*, *moderate depression {310496002}*, or *severe depression {310497006}*). If the user condition was assessed as *severe depression*, the system will select a mandatory action called *Crisis Plan*. When implemented, this action produces a dialogue in the virtual agent to inform the user about the (pre-configured) supporting contacts and tries to calm him. Additionally, the system can send an email to alert a specialist to establish a direct contact. Complementarily, if the user condition was assessed as not *severe depression*, but the obtained score in the PHQ9 still correspond to a *mild/moderate depression*, the RBS will launch some special activities in order to support the user (such as *Show a Happy/relaxing Video*).

¹ The International Health Terminology Standards Development Organization: <http://www.ihtsdo.org/snomed-ct>

CODE	ONTOLOGY	ID_SESSION	NAME	VALUE	VALUE2	DATETIME
1000001	Internal	85071hc7jbareo5tdf1vutict5	login_Completed	OK	Response from DM	2014-12-10T15:08:18.272+01:...
30000190	Internal	85071hc7jbareo5tdf1vutict5	FirstExecutionDate	OK	Initialise by the System	2014-12-10T15:08:18.278+01:...
80000450	Internal	85071hc7jbareo5tdf1vutict5	Canceled_Sessions_By_Bad_Status	0	Update value by the System	2014-12-10T15:08:18.321+01:...
80000113	Internal	85071hc7jbareo5tdf1vutict5	PMS_Device_last		Initialise by the System	2014-12-10T15:08:18.328+01:...
80000165	Internal	85071hc7jbareo5tdf1vutict5	lastReport	Not Defined	Initialise by the System	2014-12-10T15:08:18.347+01:...
80000001	Internal	85071hc7jbareo5tdf1vutict5	Session	Not Finished	not defined	2014-12-10T15:08:18.385+01:...
80000115	Internal	85071hc7jbareo5tdf1vutict5	Session_Time	513688	Update value by the System	2014-12-10T15:08:18.387+01:...
80000170	Internal	85071hc7jbareo5tdf1vutict5	Welcome_First_Time_Completed	Not Defined	Response from DM	2014-12-10T15:09:49.142+01:...
80000116	Internal	85071hc7jbareo5tdf1vutict5	DailyMoodCheck_1_Completed	3	GOOD	2014-12-10T15:10:07.146+01:...
80000119	Internal	85071hc7jbareo5tdf1vutict5	DailyMoodCheck_4_Completed	0	GOOD	2014-12-10T15:10:13.112+01:...
80000117	Internal	85071hc7jbareo5tdf1vutict5	DailyMoodCheck_2_Completed	3	GOOD	2014-12-10T15:10:17.785+01:...
80000120	Internal	85071hc7jbareo5tdf1vutict5	DailyMoodCheck_5_Completed	2	NORMAL	2014-12-10T15:10:21.319+01:...
80000118	Internal	85071hc7jbareo5tdf1vutict5	DailyMoodCheck_3_Completed	1	GOOD	2014-12-10T15:10:26.566+01:...
80000199	Internal	85071hc7jbareo5tdf1vutict5	Select_Session_Type_Selection	Long_Session	80000196	2014-12-10T15:10:46.284+01:...
80000004	Internal	85071hc7jbareo5tdf1vutict5	Mini_PHQ-9_Completed	0	GOOD	2014-12-10T15:11:05.218+01:...
80000128	Internal	85071hc7jbareo5tdf1vutict5	Speech_WavFilePath	Jgrabaciones/1_2014-12-10...	Response from DM	2014-12-10T15:11:37.452+01:...
80000130	Internal	85071hc7jbareo5tdf1vutict5	Speech_Complete	Speech_Today	Response from DM	2014-12-10T15:11:37.460+01:...
70000000	Internal	85071hc7jbareo5tdf1vutict5	Behaviour_Activation_Activities	enable	70000002	2014-12-10T15:12:06.967+01:...
70000000	Internal	85071hc7jbareo5tdf1vutict5	Behaviour_Activation_Activities	enable	70000001	2014-12-10T15:12:06.975+01:...
80000501	Internal	85071hc7jbareo5tdf1vutict5	Old_Activity_Plan	OK		2014-12-10T15:12:10.021+01:...
80000502	Internal	85071hc7jbareo5tdf1vutict5	Set_Activity_Plan_Complete	OK	70000001 70000002	2014-12-10T15:12:10.025+01:...
80000202	Internal	85071hc7jbareo5tdf1vutict5	Sleep_QuestionNumber_2	22:00		2014-12-10T15:13:06.574+01:...
80000204	Internal	85071hc7jbareo5tdf1vutict5	Sleep_QuestionNumber_4	Más de 5		2014-12-10T15:13:06.576+01:...
80000209	Internal	85071hc7jbareo5tdf1vutict5	Sleep_QuestionNumber_9	5	NORMAL	2014-12-10T15:13:06.578+01:...
80000206	Internal	85071hc7jbareo5tdf1vutict5	Sleep_QuestionNumber_6	1		2014-12-10T15:13:06.579+01:...
80000201	Internal	85071hc7jbareo5tdf1vutict5	Sleep_QuestionNumber_1	22:00		2014-12-10T15:13:06.582+01:...
80000207	Internal	85071hc7jbareo5tdf1vutict5	Sleep_QuestionNumber_7	0		2014-12-10T15:13:06.588+01:...
80000208	Internal	85071hc7jbareo5tdf1vutict5	Sleep_QuestionNumber_8	16-30		2014-12-10T15:13:06.593+01:...
80000205	Internal	85071hc7jbareo5tdf1vutict5	Sleep_QuestionNumber_5	5		2014-12-10T15:13:06.595+01:...
80000203	Internal	85071hc7jbareo5tdf1vutict5	Sleep_QuestionNumber_3	23:00		2014-12-10T15:13:06.598+01:...
80000200	Internal	85071hc7jbareo5tdf1vutict5	Sleep_Questionnaire_Complete	OK		2014-12-10T15:13:06.600+01:...
80000144	Internal	85071hc7jbareo5tdf1vutict5	Negative_Situation_Domain_Select...	¿En tus relaciones personal...	14000002	2014-12-10T15:13:36.408+01:...
80000145	Internal	85071hc7jbareo5tdf1vutict5	Negative_Situation_Inference_Sele...	Pensamientos sobre los de...	14000005	2014-12-10T15:13:43.774+01:...
80000149	Internal	85071hc7jbareo5tdf1vutict5	Negative_Situation_Text_REFUSED	Not Defined	Response from DM	2014-12-10T15:13:53.397+01:...
10000184	Internal	85071hc7jbareo5tdf1vutict5	Negative_Thought_Selection	No puedo soportar caer mal ...	12000009	2014-12-10T15:14:04.671+01:...
10000186	Internal	85071hc7jbareo5tdf1vutict5	Alternative_Thought_Selection	Sí, pero será difícil	16000004	2014-12-10T15:14:30.205+01:...
80000146	Internal	85071hc7jbareo5tdf1vutict5	Positive_Situation_Domain_Select...	¿Ha sido en casa o en relac...	14000007	2014-12-10T15:14:42.623+01:...
80000147	Internal	85071hc7jbareo5tdf1vutict5	Positive_Situation_Inference_Selec...	Pensamientos sobre mi mis...	14000010	2014-12-10T15:14:48.431+01:...
80000156	Internal	85071hc7jbareo5tdf1vutict5	Positive_Situation_Text_Complete	He sido capaz de alcanzar o...	Response from DM	2014-12-10T15:15:28.824+01:...
10000185	Internal	85071hc7jbareo5tdf1vutict5	Positive_Thought_Selection	Es posible hacer las cosas ...	13000016	2014-12-10T15:15:49.139+01:...
80000136	Internal	85071hc7jbareo5tdf1vutict5	Negative_Thoughts_Complete	OK	Response from DM	2014-12-10T15:15:49.142+01:...
80000057	Internal	85071hc7jbareo5tdf1vutict5	Beck_7	0	GOOD	2014-12-10T15:16:31.914+01:...
80000070	Internal	85071hc7jbareo5tdf1vutict5	Beck_20	0	GOOD	2014-12-10T15:16:31.920+01:...
80000050	Internal	85071hc7jbareo5tdf1vutict5	Beck_Complete	5	Leve	2014-12-10T15:16:31.922+01:...
80000066	Internal	85071hc7jbareo5tdf1vutict5	Beck_16	0	GOOD	2014-12-10T15:16:31.926+01:...
80000067	Internal	85071hc7jbareo5tdf1vutict5	Beck_17	0	GOOD	2014-12-10T15:16:31.929+01:...
80000061	Internal	85071hc7jbareo5tdf1vutict5	Beck_1	1	GOOD	2014-12-10T15:16:31.932+01:...

Figure 1. Data stored in the model of the user. The data is updated by the cognitive layer, which collects objective information such as questionnaire's scores, and the inference of subjective evaluations.

3.2. Session Planner Layer

In this layer, all the inferred tasks in the cognitive layer must be managed and selected for inclusion in the interactive session. All the activities inferred in the cognitive layer are based on clinical expert knowledge, but the session planner filter these activities and arrange them on a specific order based on (1) a set of predefined clinical requirements; and on (2) the historical user data (included in the user model). The set of clinical requirements are predefined according to clinicians' preferences or protocol to offer different activities used to detect and prevent depression or suicidal ideation. As we show in *code 1*, each activity defined by the clinical expert must contain: a priority; its dependencies (some tasks could be blocked or enabled by other tasks); and periodicity (the minimum and maximum number of executions during a week). These settings can produce different types of sessions according to the characteristics of the addressed users or according to the preferred prevention protocol from the domain experts. As evaluated in (Bresó, et al., 2015), different settings produce different contents in the sessions and the interaction can be more flexible or more restrictive in terms of variability and adaptability. The second set of settings that influence the specific content of each session is the information stored in the user's model that includes past actions executed in previous sessions. The model of the user stores all these data such as activities blocked by the user, the acceptance or not of the proposed activities, and the number of times that the user has performed the proposed activities. All this information is taken into account to plan the contents of the actual session.

```
<?xml version="1.0" encoding="UTF-8"?>
...
<item>
  <description>
    <code>10000203</code>
```

```

        <snomed>NO</snomed>
        <name>Hopeless_Beck_Questionnaire</name>
    </description>
    <enable>YES</enable>
    <typeOfItem>OPTIONAL</typeOfItem>
    <priority>20</priority>
    <minFrec>2</minFrec>
    <maxFrec>3</maxFrec>
    <codeOfDataInDB>80000050</codeOfDataInDB> <!--Beck_Complete-->
    <repeatable>NO</repeatable>
    <constrains>
        <constrain>
            <constrainDescription>
                <code>10000011</code>
                <snomed>NO</snomed>
                <name>Welcome</name>
            </constrainDescription>
            <constrainType>BEFORE</constrainType>
        </constrain>
        <constrain>
            <constrainDescription>
                <code>10000200</code>
                <snomed>NO</snomed>
                <name>Introduce_PHQ9</name>
            </constrainDescription>
            <constrainType>BEFORE</constrainType>
        </constrain>
        <constrain>
            <constrainDescription>
                <code>10000121</code>
                <snomed>NO</snomed>
                <name>Farewell</name>
            </constrainDescription>
            <constrainType>AFTER</constrainType>
        </constrain>
    </constrains>
</item>
...
</xml>

```

Code 1: Example of the clinical requirements for the task *Hopeless Beck Questionnaire* established in the xml configuration file. Clinicians may define constrains such as priority, dependence, and periodicity.

3.3. Emotional Layer

Equally important to the contents of the sessions is how these contents are transmitted to the users in order to make them more receptive and promote an effective execution of the suggested activities and the disclosing of personal information such as thoughts and feelings. The use of empathic agents as the virtual peer of the user has been recently increased in applications to support change behavior and in cyber-psychology (Bresó, et al., 2015). One relevant aspect that these interactive tools should implement is the generation of a coherent and adequate emotional behavior while interacting with the users. As presented in different studies (Bickmore & Gruber, 2010) (Lisetti & L., 2008) the conveying of *empathic reactions* from the virtual agent to the information provided by the user is a key characteristic to increase the acceptability of these virtual peers in the users.

The emotional layer of our system is the component responsible to generate the *adequate* emotions in the virtual agent to produce an *empathic feedback* based on the user's input at each cycle of interaction. This component is based on an existent computational architecture of emotions known as FATiMA (Dias, et al., 2014), which in turn is based on the cognitive appraisal theory of emotions (Scherer, et al., 2001). The key assumption of the appraisal theory of emotions is that the events produced in a person's environment are constantly evaluated by the individual. This cognitive evaluation (or appraisal) process

leads to an emotional response (according to the relevance that the event has for the person), which in turn generates a specific behavior to cope with the appraised events. Based on this theory, the FAtiMA architecture offers a generic appraisal framework where a set of different *basic* emotions and the coping behaviors can be generated according to a set of goals, preferences and action tendencies predefined for a virtual agent.

Based on the context of application, we have scripted all these goals, preferences and action tendencies for our virtual agent. The events produced during the interaction with the user –i.e. the knowledge extracted in the cognitive layer– generates the specific emotional behavior of the virtual agent. Most of the goals defined in the virtual agent are close related with the wellbeing detected in the user for the agent to produce an empathic reaction to the detected user's state. Taking again the example of the user responses to the PHQ-9 questionnaire, if the score obtained indicates that the depression level is quite high, the virtual agent appraises this result as a highly undesirable event for the user's optimal wellbeing condition, which generates a negative emotion. This negative emotion is displayed through the verbal (specific dialogue) and non-verbal (specific facial expression) communication during the interaction with the user to convey a sense of understanding of the situation.

Moreover, we have extended the FAtiMA architecture to generate a more user-adapted style of interaction in the virtual agent according to the historical information taken from the model of the user (Martínez-Miranda, et al., 2014). This extension includes a re-appraisal process of the detected event in order to alter the initial significance of that event and that can produce a more personalized feedback to the user. Following with the PHQ-9 questionnaire example, if the obtained score indicates a high level of depression, the re-appraisal component can consult the results obtained in the PHQ-9 questionnaire during previous sessions (stored in the model of the user) and check whether the current result shows a positive tendency in the user's condition taking into account the previous results. If a positive tendency is found, the original event would be reappraised as "not much undesirable" to the user (though the current PHQ-9 score is still not the optimal). This reappraisal can change the emotional state or the emotion's intensity in the virtual agent, which is reflected in the specific feedback provided to the user.

3.4. Visualization Layer

All the contents of a session produced by the cognitive and the session planning layers, as well as the representation of the different styles of interaction formed by the different emotions produced in the emotional layer needs to be visually represented in the GUI which contains the embodied virtual agent. The virtual agent has been developed as a talking head using a commercial tool called FaceShift (www.faceshift.com). This software captures the facial movements of a person in real time using a xbox 360 Kinect camera for the creation of a set of animated avatars. The results present a good level of realism and allow the representation of different emotional expressions. A set of video recordings was performed representing different facial expressions that are linked in real time with the corresponding emotion generated by the emotional layer. This allows the visualization of the empathic feedback provided by the virtual agent at each cycle of interaction. Realism in the appearance and behaviour in a virtual agent are key issues to achieve good levels of acceptability in the target users.



Figure 2: Recording a video using the Faceshift tool, which uses a kinect xbox 360 camera



Figure 3: The graphical user interface with the virtual agent situated at the centre of the screen. This virtual agent conducts the session, providing questionnaires, recommendations or CBT recommended tasks to the user.

4. Case of study

The components presented in the previous section have been integrated in a complete application evaluated in the context of a Spanish regional project entitled “*PrevenDep: Computational system to help in the prevention of depression and suicide*”. The developed application has been adapted to provide the necessary support to university students for detecting and preventing depression and suicide. University students are exposed to challenging and competitive environment in which they must to overcome new situations, stress, academic challenges, or social relationships. PrevenDep provides an empathic virtual agent to support them by offering user-adapted activities based on CBT interventions (such as a *negative thoughts questionnaire*, which tries to identify and transform negative thoughts of the users into an alternative positive perspective; or the *configuration of an activity plan*, which agrees with the users few routine activities that should be performed during the week and that they must report).

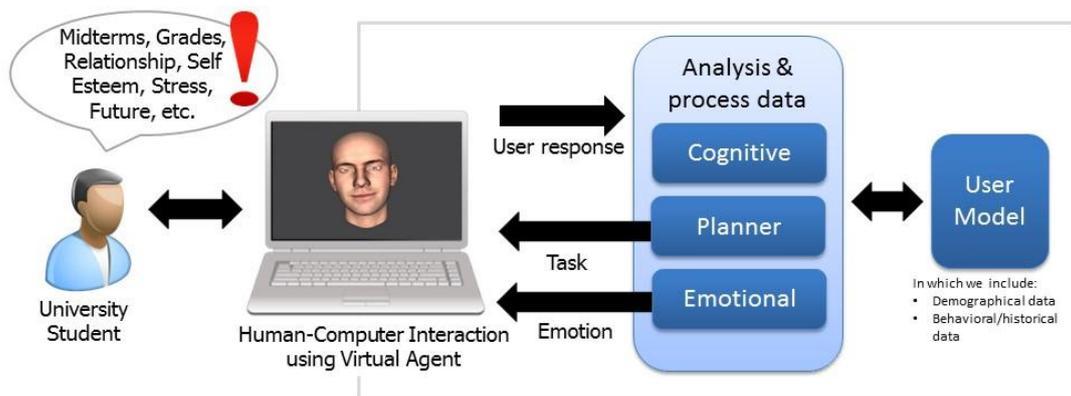


Figure 4: Overall architecture of the system in the PrevenDep project. An Empathic Virtual Agent + GUI form the visualization layer of the interactive sessions with a user. The core of the system is composed by the other three components: the cognitive,

the session planning, and the emotional layers. The functionalities of all the layers are based on a dynamic user model, in which the system stores and updates all the information of the user for an adequate adaptation and personalization of the interactive sessions.

4.1. Objective

The aim of the PrevenDep testing was to run a clinical trial with a set of potential real users from the focus group composed by students from different academic institutions in order to study its efficacy. Nevertheless, attending security and ethical reasons, before the use of the system with the target users, we have performed an evaluation with a set of professors and researchers from different graduate and postgraduate institutions in Spain and Mexico. This evaluation has provided us a valuable opinion from an expert group, which closely interact with the potential final users. This evaluation has been focused to obtain feedback about the acceptability and usability levels of the system.

4.2. Participants

A recruitment screening survey and consent form were posted by mail to approximately 100 professors, researchers and students involved in their final degree projects from four different locations: *Universitat Politècnica de Valencia*², *Universitat de Valencia*³, *Universitat Jaume I*⁴, and *CICESE-UT3*⁵.

4.3. Evaluation

To carry out the evaluation, an explanatory website was developed where a set of videos containing several executions of different interactive sessions produced along several days showing the different contents and reactions from the virtual agent according to the user inputs. The visualization of the videos was as a prerequisite to complete the evaluation questionnaire. In order to measure the usability and acceptability of the system, we designed a questionnaire with a total of 22 questions (q1-q22), grouped into 6 different sections: 1) **demographic information**; 2) **system usability**; 3) **user acceptability of the contents of the session** that evaluates how appropriate the user considers the activities proposed by the virtual agent; 4) **user acceptability of the virtual agent** that evaluates the appearance and behavior of the virtual agent; plus two free text sections to write any feedback the participants wants to provide about the 5) session contents and about the 6) aspect and behavior of the VA (see Table 1).

Section	Question	Text
1 Demographic information.	q1	Please provide your age
	q2	Please provide your gender
	q3	Please provide the institution where you usually develop your work
2 Usability of the system	q4	I think the target users will like to visit frequently the PrevenDep system
	q5	I think that target users will find the PrevenDep system unnecessarily complex to use
	q6	I think that the PrevenDep system is easy to use.
	q7	I think that the target users will need the support of an expert to use the PrevenDep system
	q8	I think that target users will find the various possibilities of PrevenDep fairly well integrated in the system
	q9	I think that the user would find too much inconsistency in the PrevenDep system
	q10	I think that most of the potential users will learn to use the PrevenDep system in a short time

² www.upv.es

³ www.uv.es

⁴ www.uji.es

⁵ www.cicese.edu.mx

		q11	I found PrevenDep system very cumbersome to use.
		q12	I think that the target users will be very confident when using the PrevenDep system
		q13	I think there is necessary to learn some previous things before using adequately the PrevenDep system
3	User acceptability of the session contents.	q14	I think the length of the sessions was adequate, allowing the user to complete the sessions on daily basis.
		q15	I think that the variability of the sessions was adequate, which motivate the user to complete the whole sessions on daily basis.
		q16	I think that the content of the sessions was adequate and would motivate the user to complete the whole sessions on daily basis.
4	User acceptability of the Virtual Agent.	q17	The behavior of the virtual agent during the sessions seems aloof and distant.
		q18	The virtual agent inspires trust.
		q19	The virtual agent behaves emotionally stable.
		q20	The behavior of the virtual agent motivates the daily use of the PrevenDep system.
5	Additional comments I: regarding the contents of the session	q21	Please discuss any unpleasant aspect about the contents of the daily sessions.
6	Additional comments II: regarding the interaction with the Virtual Agent	q22	Please discuss any unpleasant aspect in the appearance and / or behavior of the virtual agent.

Table 1: Questionnaire used in the evaluation containing twenty-two questions grouped in 6 sections.

The first section (q1-q3) collects the demographic data (age, gender and workplace) through three categorical questions. The second section of the questionnaire is composed by ten questions (q4-q13), which correspond to the standardized System Usability Scale Questionnaire (SUS) (Brooke, 1996). The SUS questionnaire was created by the Digital Equipment Corporation and it is mainly focused on the assessment of perceived usability (subjective aspects of usability) on any system. In the third section, three questions (q14-q16) are used to determine the acceptability of the most important features of the session planner: the adaptability (contents and length of the session) and variability (different contents between sessions). The questions in section number four (q17-q20) are used to identify the acceptability of the users regarding the appearance and behavior of the VA. Finally, in sections 5 (q21) and 6 (q22), two free text questions are used to collect any comments regarding the session planner and the VA.

The answers of each item in sections 2, 3 and 4 are defined on a Likert scale with values between 0 and 4, representing the degree of agreement with the statement as shown in Table 2.

Code	Response
0	Strongly Disagree
1	Disagree
2	Neither Agree nor Disagree
3	Agree
4	Strongly Agree

Table 2: Likert codes used in sections 2, 3 and 4

5. Results

The questionnaire was online during 15 days and a total of 60 participants accessed to the evaluation. From 60 participants, the 93.6% fully complete the evaluation questionnaire. Specifically, the participation in the first 20 questions was 96.4% and in the two last questions (q21 and q22) was 65.8%.

5.1. Quantitative findings

Regarding the answers to the demographic questions we obtained that the gender distribution included 26 (43.3%) women, 33 (55%) men, and one participant omitted this information (1.7%). Using the information of the gender, we checked the A.L.Baylor hypothesis (Baylor, 2009) that stated that a virtual agent is usually rated higher by those participants with the same gender modeled in the VA. Nevertheless, in our evaluation we did not find any significant difference in the acceptance of virtual agent based on the gender of the participant.

The age distribution was 2 people (3.3%) under 25; 18 participants (30%) between 26 and 30; 15 people (25%) between 31 and 35; 12 participants (20%) from 31 to 35; 9 participants (15%) between 31 and 35; and 4 people (6.7%) over 46. Regarding the center of belonging, the majority was the UPV with 28 participants (46.7%); UJI with 13 participants (21.7%); 10 participants (16.7%) from UV; 5 individuals (8.3%) from CICESE; and 4 participants (6.7%) from others institutions. The normalized results are showed in the following figures.

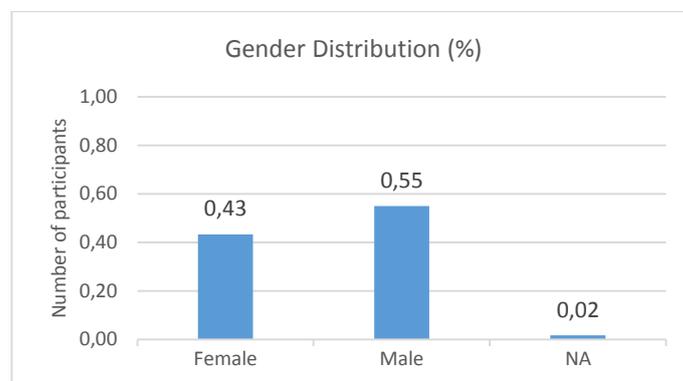


Figure 5: Participants gender distribution obtained from question q1 of the demographic section, in the evaluation questionnaire.

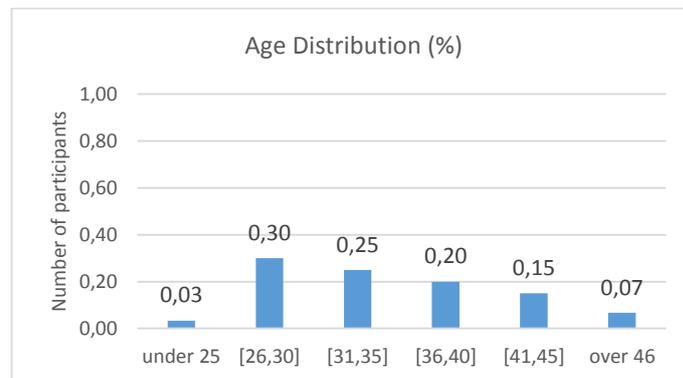


Figure 6: Participants age distribution obtained from question q2 of the demographic section, in the evaluation questionnaire.

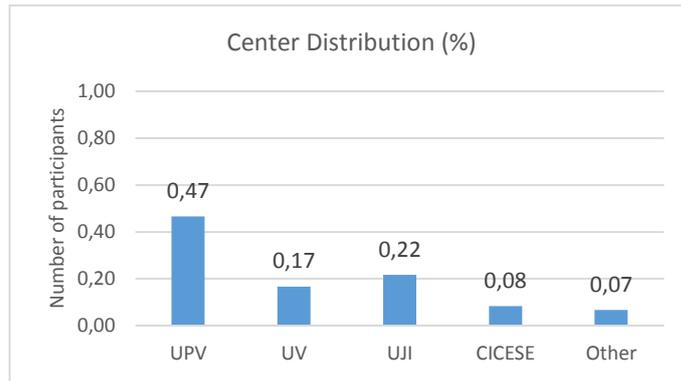


Figure 7: Participants center distribution obtained from question q3 of the demographic section, in the evaluation questionnaire.

The results from the usability questionnaire (section 2 of the evaluation questionnaire) are summarized in Table 3 showing for each question (rows) the response rate and the number of answers (columns). The questions and the meaning of the answers are those already introduced in Tables 1 and 2 respectively. The last column indicates the weighted average.

High rate responses to questions with even numbers (i.e. q4, q6, q8, q10 and q12) represents a positive outcome of the feature evaluated in the system while low rates responses to questions with odd numbers (i.e q5, q7, q9 and q11) indicate a positive outcome on the evaluated feature. A positive normalization was performed over the questions with odd numbers by reversing the results and a threshold value (=2) was defined to graphically represent the obtained results. We can see in Figure 8 that the mean of the obtained responses passed the defined threshold.

	0	1	2	3	4	Weighted average
q4	0% (0)	16% (9)	24% (14)	40% (23)	21% (12)	2.66 (58)
q5	46% (26)	42% (24)	5% (3)	5% (3)	2% (1)	0.75 (57)
q6	0% (0)	4% (2)	4% (2)	53% (30)	40% (23)	3.30 (57)
q7	35% (20)	46% (26)	11% (6)	9% (5)	0% (0)	0.93 (57)
q8	0% (0)	4% (2)	30% (17)	41% (23)	25% (14)	2.88 (56)
q9	39% (22)	39% (22)	18% (10)	5% (3)	0% (0)	0.89 (57)
q10	0% (0)	0% (0)	5% (3)	37% (21)	58% (33)	3.53 (57)
q11	14% (8)	34% (20)	38% (22)	9% (5)	5% (3)	1.57 (58)
q12	0% (0)	11% (6)	33% (19)	35% (20)	21% (12)	2.67 (57)
q13	43% (25)	47% (27)	7% (4)	2% (1)	2% (1)	0.72 (58)

Table 4: Results of the System Usability Scale Questionnaire (SUS) belonging to the second section of the questionnaire. Answers are coded using a Likert scale defined in the Table 2. The last column shows the weighted average responses (value from 0 to 4) and the number of participants.

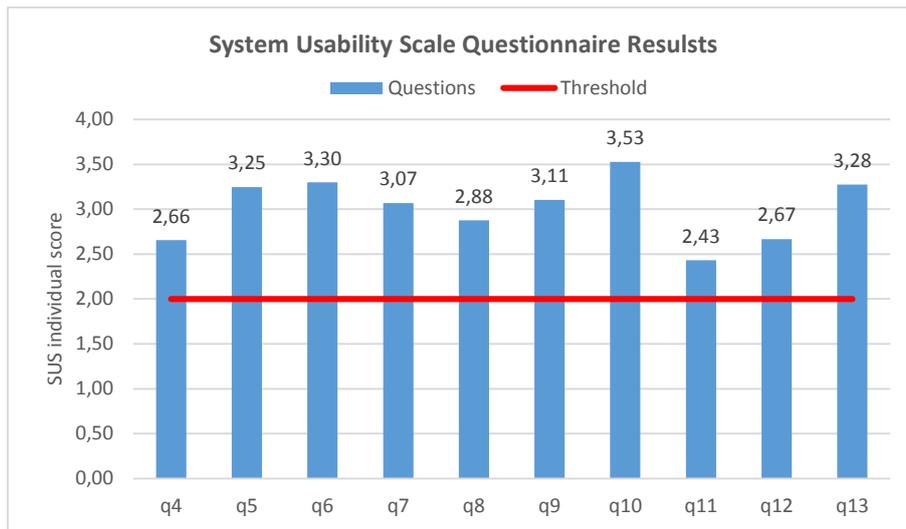


Figure 8: Graphical representation of the results of the questionnaire System Usability Scale (SUS), corresponding to the q4-q13 questions. Results from the questions with odd values (q5, q7, q9 and q11) were inverted (positive normalization) to facilitate the interpretation of the obtained results. All the results exceed the threshold meaning a good assessment in the perceived usability of the system

The SUS questionnaire proposes a specific method to evaluate the obtained results. To calculate the score of SUS, we must separate the positive questions –1 (q4), 3 (q6), 5 (q8), 7 (q10) and 9 (q12)– from the negative questions –2 (q5), 4 (q7), 6 (q9), 8 (q11) and 10 (q13)–. For the positive questions, we updated their values from the obtained value minus 1. For negative questions, we used 5 minus the original value. Finally, we performed an arithmetic sum of all the updated values and multiplied them by 2.5 in order to normalize the final result between 0 and 100. Based on more than 500 evaluations, Jeff Sauro (Sauro, 2011) argues that the average score to assess the usability of a system should be a value of 68. A higher value means a usability level better than the average. Figure 9 presents the average result obtained from the collected responses corresponding to a value of 75.7 (gray line), which is above the average proposed by Jeff Sauro (orange line). Therefore we can conclude that participants in our evaluation rated the system with a high usability feature.

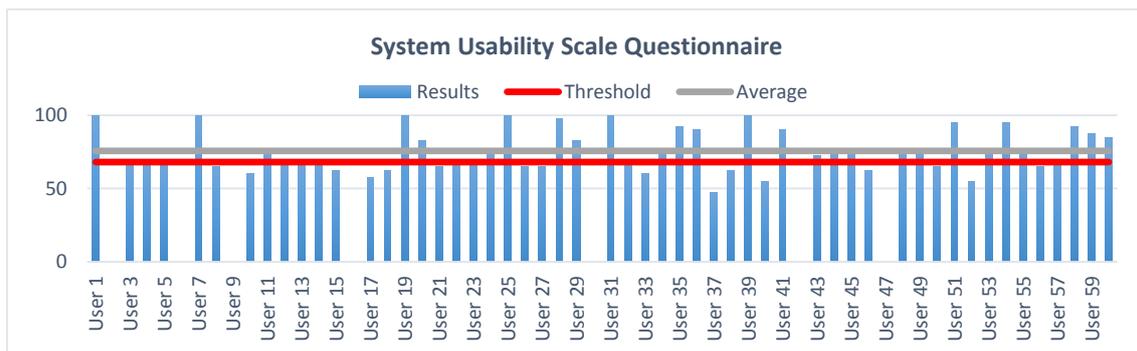


Figure 9: Graphical representation of the results obtained in section 2 of the questionnaire (perceived usability of the system). The results were obtained following the method proposed in [26]. Most of the participants correctly completed the questionnaire and the average of the results (gray line) is located above the average score proposed by Jeff Sauro (orange line).

The obtained results from the questions of section three collect the participants’ feedback regarding the suitability and variability of the sessions’ contents. The question 14 (q14) is focused in the assessment of the session’s length. The produced length of a session is highly important taking into account that the system should adapt the length of each session according to the status (e.g. mood) detected in the user and at the same time meeting the clinician’s requirements (i.e. the scheduling of certain activities a minimum number of times per week). Complementarily, question 15 (q15) is focused on the perceived variability in the contents of the sessions. A strategy to maintain motivated the users with the use of the

system is that the sessions do not offer a monotonous or repetitive content. Finally, question 16 (q16) is focused to evaluate the perceived usefulness of the content of the sessions, i.e. how much helpful are perceived the activities or recommendations offered by the system. The results of this section are summarized in Table 5 and graphically presented in Figure 10. Similarly to the obtained results from section 2, we can see that the responses of the participants indicate positive results, particularly the assessment of the length of the session (q14).

	0	1	2	3	4	Weighted average
q14	0% (0)	2% (1)	9% (5)	55% (32)	34% (20)	3.22 (58)
q15	0% (0)	14% (8)	31% (18)	31% (18)	24% (14)	2.66 (58)
q16	0% (0)	0% (0)	34% (20)	34% (20)	31% (18)	2.97 (58)

Table 5: Obtained results in section three of the questionnaire: assessment of the sessions' contents (q14, q15 and q16)

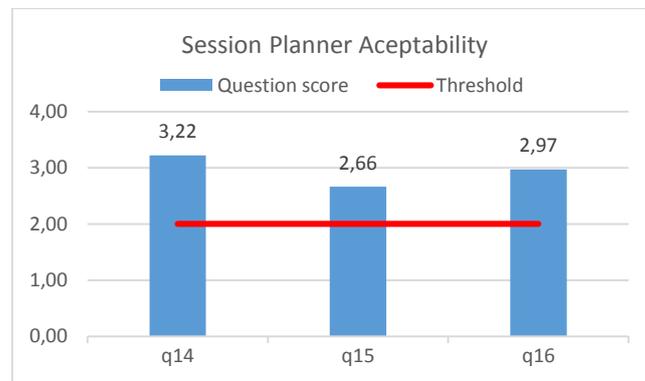


Figure 10: Graphical representation of the evaluation of the sessions' contents (q14, q15 and q16)

The objective of the questions in the section fourth of the questionnaire was very similar to the questions of the third section, but now focused in the evaluation of the behavior and appearance of the virtual agent. The building of a good relationship between a virtual agent and the user is essential to promote a continuous use of the system and a key ingredient to establish a therapeutic alliance that benefit the user. The four questions of section three of the questionnaire looked to obtain the feedback from the participants regarding the emotional behavior of the virtual agent (q17, q18 and q19) and how much the virtual agent contributes to the motivation of the user to execute the system every day (q19). The obtained results show a positive assessment in each of the features represented in the four questions (see Table 6 and Figure 11). The 66% of the participants were disagreed or strongly disagreed with the sentence that defines the virtual agents as "distant and aloof" (q17). Complementarily, the 57% of the participants perceived the virtual agent as trustworthy (q18). For the 82% of the participants the virtual agent was perceived as "emotionally stable" (q19), and the 48% consider that the virtual agent is a decisive factor that motivate the use of the system (q20).

	0	1	2	3	4	Weighted average
q17	26% (15)	36% (21)	14% (8)	24% (14)	0% (0)	1,36 (58)
q18	0% (0)	16% (9)	28% (16)	36% (21)	21% (12)	2,62 (58)
q19	0% (0)	5% (3)	12% (7)	41% (24)	41% (24)	3,19 (58)
q20	2% (1)	10% (6)	40% (23)	26% (15)	22% (13)	2,57 (58)

Table 6: Obtained results related with the evaluation of the virtual agent (q17, q18, q19 and q20).

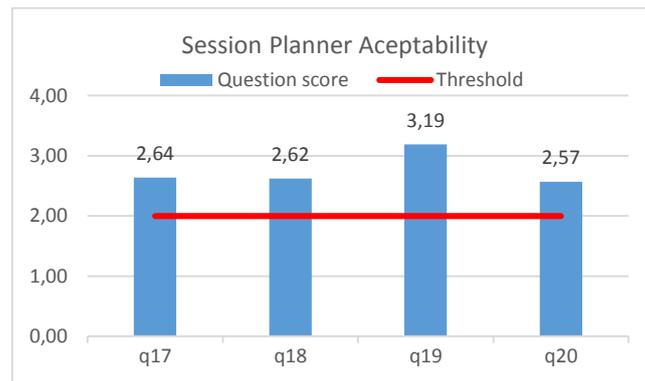


Figure 11: Graphical representation of results obtained in the fourth section of the questionnaire (q17, q18, q19 and q20). The value of the question q17 had been normalized to positive version.

5.2. Qualitative findings

While the first 20 questions of the questionnaire have been designed to collect the perceived usefulness from the participants regarding the different features of the system, the design of the last two questions were thought to collect specific feedback that can be used to improve the functionality of the session planner (q21) and of the virtual agent (q22). Many comments were positive but also many of them addressed particular issues that can be improved before the execution of the pre-clinical trial of the system. Some of the most representative comments to improve the system related with the session contents are transcribed below:

Participant P51: *“Some questionnaires are too long and take up too much space. It would be convenient to divide them into several screens to prevent the rejection of the user. Additionally, the font size in the questions should be bigger to ensure readability and attractiveness of the appearance of the interface”.*

Participant P50: *“I like the content but I miss questions related with the university environment. This system is focused on students, so why not ask things directly related with the university e.g. the exams?, trend of success or failure of the student?, number of registered credits (to know the pressure bearing)?, do you live alone?, combines study with work?, what are your study habits?, is everything done at the last minute?. I think that this kind of information is needed to better identify a risk of suicide”.*

Participant P49: *“The sessions and their content seem appropriate, easy to understand, useful and necessary to assess important aspects related to depression”.*

Participant P44: *“Sometimes I have the feeling that during a session there is more content related to assessment (questionnaires) that content related with therapeutic intervention (relaxing activities, more information, etc.)”.*

Participant P36: *“Regarding the content of the sessions, I think it would be positive to include at the end of each of session, a question that explore the perceived utility”.*

Participant P21: *“The appearance of the application is unattractive, which may have implications for adherence to the system. Several suggestions in this regard: It would be good that the questions and answers are read with voice; the virtual agent must be always present in the screen.”*

Perhaps some questionnaires are too extensive. Finally, I recommend a control of achievements, for example, having visible information on progress in therapy".

Participant P4: "In general, the content is adequate but the questionnaires were very long and boring, could do a little more interactive? I prefer that the virtual agent is placed always in the corner, interacting with the user. The user must interact using verbal communication".

From the collected comments, we can identify that some of them are more focused on how the contents of the session are presented through the GUI (P4, P21, P37 and P51) and others in the contents of the sessions themselves (P36, P44, P49 and P50). From the first group we can use the provided suggestions to get a better GUI through the improvement of the controls (buttons, combo boxes, font size). As the current state of the technology does not allow a full natural language interaction free of errors, we decided to maintain the input from the user through the buttons and other controls such as combo boxes included in the GUI and not to receive the input from the user through natural language processing. From the second group of suggestions we can observe that more context-based questions (i.e. from the university context) needs to be included, as well as some question(s) about the perceived usefulness from the target users regarding the contents offered by the system. Moreover, we also observed that the inclusion of more therapy-based activities has been suggested.

Regarding the comments on the appearance and behavior of the virtual agent (q22), we also obtained many useful comments. In general, many users acknowledge a suitable behavior in the virtual agent and most of the comments were related with the improvement of its appearance. The transcription of some of the received comments is the following:

Participant P60: "In general, the virtual agent behavior seems appropriate. The face looks nice and the expressive voice gives me confidence. The disadvantage that I see is some facial expressions (sometimes too expressive, overplayed, and there are some tics in the eyebrow)".

Participant P53: "Displaying only the head on a black background does not seem to be a nice picture. I do not understand why the avatar is hairless. His tone of voice sounds negative. Do you think about the possibility of choosing between a male and female virtual agent? And the environment, why do not show a clinical site (as a therapy room with a table and chairs and perhaps a window where you could visualize a landscape outside)?, I think that would increase the effectiveness of the tool".

Participant P49: "It is very good but can be improved a bit (need a shave and cleaning of teeth). I really like movements. It is the most convincing I've seen".

Participant P41: "The virtual agent can be improved. I think it would be better if it would be more protagonist during the interaction (more talking and less read). I would like that it have hair, he was a little older and it should have shoulders (with a shirt or similar). I don't like his teeth. But overall I think it looks very real, I like the way it moves. Eyebrow movement is very good. But you have to work a bit, you have many possibilities".

Participant P37: "Do you inspire more confidence with a bald man? Although I like it, I suggest that you consider the option that each user can configure his virtual agent".

Participant P21: "The virtual agent has a warm and calm voice. However, I consider the appearance can be improved".

Almost all of the comments suggested improvements to the appearance of the virtual agent more than to the perceived behavior. Although the Faceshift software allows a fast prototyping of a realistic virtual

agent, a deeper programming work is required to customize the appearance of the pre-defined virtual agent (i.e. get a set of virtual agents with different gender and age, adding a half or full body or even adding hair and clothes). The further improvement of the suggested features in the appearance of the virtual agent would even positively influence the perception of a more suitable behavior during the interactive sessions, which in turn could help to optimize the effectiveness of the system.

6. Conclusions and Future Work

The design and development of tools such as the current system presented, becomes a prime target from a clinical point of view, given the enormous prevalence of depression and suicidal ideation, especially in young people. In fact, psychotherapy research needs to broaden in terms of adoption of large-scale public health strategies and treatments that can be applied to more patients in a simpler and cost-effective way (Emmelkamp, et al., 2014) and ICT-based tools to prevent depression or other disorders might be one promising way forward (Kazdin & Blase, 2011).

The work described in this paper presents and evaluates an interactive system aimed to support university students in the prevention of depression and suicide. Both, the contents of the interactive sessions and the different styles of interaction (represented through a conversational virtual agent) are dynamically adapted based on the state inferred in the user.

The functionalities of each component of the system ensure a continuous user-adapted intervention, a key aspect to maximize the benefits that this type of interventions can produce in people with problems related to mental health. In a previous study we perform a functional evaluation of the components through a set of simulations (Bresó, et al., 2015). In this study we present an initial evaluation about the perceived acceptability and usability involving 60 participants who answered a questionnaire after watching a set of pre-recorded videos containing different interactive sessions. The participants rated the usability of the system as highly positive (with a result 75.7 out of 100). The results obtained regarding the level of acceptability concerning the sessions' management (content, duration and variability) have also been positive, obtaining a score of 2.95 out of 4. Regarding the evaluation of the appearance and behavior of the virtual agent –as the main interface between the user and the system–, the participants also valued positively this component (getting a score of 2.75 out of 4). We can summarize these acceptability results as a percentage of 70.9% reflecting that the participants have rated a high level of acceptability of the framework. The results obtained from this pre-evaluation are good enough to update and improve the system following all these recommendations and conduct an initial clinical trial with a set of potential real users, which is part of the future work.

Additionally, we also plan to include some of the improvements identified by the participants, especially those suggestions to get a better GUI of the system. Once implemented, we will develop a pre-clinical trial enrolling a set of University students that fulfill some criteria to be selected.

Acknowledgments

The work presented in this manuscript has been partially funded by the *Consellería de Sanidad of Generalitat Valenciana*, in the research project entitled "*Sistema computacional de ayuda a la prevención de episodios de depresión y suicidio –PREVENDEP*". We thank the company Faceshift (www.faceshift.com) for providing their software to perform facial motion capture in order to develop the talking head that represent our empathic virtual agent.

References

- Arning, K. & Ziefle, M., 2007. Understanding age differences in PDA acceptance and performance.. *Computers in Human Behavior*, 6(23), pp. 2904-2927.
- Baylor, A., 2009. Promoting motivation with virtual agents and avatars: role of visual presence and appearance. *Philos Trans R Soc Lond B Biol Sci*, 1535(364), p. 3559–3565.

- Bickmore, T. W. & Gruber, A., 2010. Relational agents in clinical psychiatry. *Harvard review of psychiatry*, 18(2), pp. 119-130.
- Bole Williams, B., 2013. *Suicide Prevention*, Rowan University: s.n.
- Bresó, A., Martínez-Miranda, J. & García-Gómez, J., 2015. A Novel Approach to Improve the Planning of Adaptive and Interactive Sessions for the treatment of Major Depression [under review]. *International Journal of Human-Computer Studies*, Volume under revision.
- Bresó, A., Martínez-Miranda, J. & García-Gómez, J., 2015. Intelligent Virtual Agents Applied in Mental Health Care: A Survey [under review]. *Journal of Medical Internet Research (JMIR)*.
- Brooke, J., 1996. SUS: A quick and dirty usability scale.. In: *Usability Evaluation in Industry*. London: Taylor & Francis: P.W. Jordan, B. Thomas, B.A. Weerdmeester & I.L. McClelland (Eds.), pp. 189-194.
- Bryan, C. & Albright, G., 2013. *A Simulation Utilizing Virtual Humans to Prepare Primary Care Providers to Conduct Substance Use and Mental Health Screening and Brief Intervention*, s.l.: s.n.
- Capuano, N., D'Aniello, G., Gaeta, A. & Miranda, S., 2015. A personality based adaptive approach for information systems. *Computers in Human Behavior*, Issue 44, pp. 156-165.
- Corrieri, S. et al., 2014. School-based prevention programs for depression and anxiety in adolescence: A systematic review. *Health promotion international*, 29(3), pp. 427-441.
- Cuijpers, P., Beekman, A. & Reynolds, C., 2012. Preventing depression: A global priority. *Journal of the American Medical Association*. <http://dx.doi.org/10.1001/jama.2012.271>, 307(10), pp. 1033-1034.
- Darves, C., Oviatt, S. & Coulston, R., 2002. Adaptation of users' spoken dialogue patterns in a conversational interface. In *Proceedings of the 7th International Conference on Spoken Language Processing*, Volume 1, pp. 561-564.
- Dias, J., Mascarenhas, S. & Paiva, A., 2014. FAtiMA Modular: Towards an Agent Architecture with a Generic Appraisal Framework. In: *Emotion Modeling: Towards Pragmatic Computational Models of Affective Processes*. s.l.:Bosse T., Broekens J., Dias J., van der Zwaan J. (Eds.), pp. 44-56.
- Dix, A., Finlay, J., Abowd, G. & Beale, R., 1993. *Human-Computer Interaction*. s.l.:Prentice Hall.
- Emmelkamp, P. et al., 2014. Advancing psychotherapy and evidence-based psychological interventions. *International Journal of Methods in Psychiatric Research*, 23(S1), pp. 58-91.
- Fischer, G., 2001. User Modeling in Human-Computer Interaction. *User Modeling and User-Adapted Interaction*, pp. 65-68.
- Griffiths, K. & Christensen, H., 2007. Internet-based mental health programs: A powerful tool in the rural medical kit. *Aust. Journal Rural Health*, Volume 15, pp. 81-87.
- Haas, A. P., Hendin, H. & Mann, J. J., 2003. Suicide in college students. *American Behavioral Scientist*, 46(9), pp. 1224-1240.
- Horgan, A., McCarthy, G. & Sweeney, J., 2013. An Evaluation of an Online Peer Support Forum for University Students With Depressive Symptoms. *Archives of Psychiatric Nursing*, 27(2), pp. 84-89.
- Joffe, P., 2008. An empirically supported program to prevent suicide in a college student population. *Suicide and Life-Threatening Behavior*, 38(1), pp. 87-103.
- Johnson, A. & Taatgen, N., 2005. User Modeling. In: *Handbook of human factors in Web design*. s.l.:Lawrence Erlbaum Associates, p. 424-439.
- Kazdin, A. E. & Blase, S. L., 2011. Rebooting psychotherapy research and practice to reduce the burden of mental illness. *Perspectives on Psychological Science*. <http://dx.doi.org/10.1177/1745691610393527>, 6(1), pp. 21-37.
- Kirsch, D. J. et al., 2014. Population-Based Initiatives in College Mental Health: Students Helping Students to Overcome Obstacles. *Current psychiatry reports*, 16(12), pp. 1-8.
- Lee, s., 2010. An ecological understanding of youth suicide in South Korea. *School Psychology International*, 5(31), p. 531.
- Lisetti & L., C., 2008. Embodied conversational agents for psychotherapy. In *CHI 2008 Workshop on Technology in Mental Health*.
- Manning, L., Manolya, K. & Tarashankar, R., 2012. eCounsellor: An Avatar for Student Exam Stress Management. *Proceedings of the 3rd International Conference on Information Management and Evaluation*, pp. 185-193.
- Martínez-Miranda, J., Bresó, A. & García-Gómez, J. M., 2014. Modelling Two Emotion Regulation Strategies as Key Features of Therapeutic Empathy. In: *Emotion Modeling: Towards Pragmatic Computational Models of Affective Processes*. s.l.:En Bosse T., Broekens J., Dias J. y van der Zwaan J. (Eds.) Springer, pp. 115-133.
- Organization, W. H., 2014. *Preventing suicide: A global imperative*. ISBN: 978 92 4 156477 9: s.n.
- Pelkonen, M., Marttunen, M. & Henriksson, M., 2005. Suicidality in adjustment disorder--clinical characteristics of adolescent outpatients. *European Child and Adolescent Psychiatry*, 14(3), pp. 174-80.
- Preece, Rogers, J. Y., Sharp, H. & Benyon, D., 1994. *Human-Computer Interaction*. s.l.:Addison-Wesley..
- Rajamanickam, M. R., 2011. Designing adaptive interfaces for children : a preliminary study on the effect of age and gender on children's interaction in the context of dialoguing with computers. *Master of Science. Program Computer Science. University of British Columbia*.
- Richards, D., 2009. Features and benefits of online counselling: Trinity College online mental health community. *British Journal of Guidance & Counselling*, 3(37), pp. 231-242.

- Sauro, J., 2011. *A Practical Guide to the System Usability Scale: Background, Benchmarks & Best Practices*. Denver: ISBN: 978-1461062707.
- Scherer, K. R., Schorr, A. & Johnstone, T., 2001. *Appraisal Processes in Emotion: Theory, Methods, Research*. Oxford University Press: s.n.
- Schwartz, A. J., 2006. College student suicide in the United States: 1990–1991 through 2003–2004. *Journal of American College Health*, 54(6), p. 341–352.
- Tarashankar, R., Manning, L. & Manolya, K., 2012. *ESCAP: Towards the Design of an AI Architecture for a Virtual Counselor to Tackle Students' Exam Stress*. Hawaii , s.n.
- Torre Martí, M., 2013. *PROTOCOLO PARA LA DETECCIÓN Y MANEJO INICIAL DE LA IDEACIÓN SUICIDA*, https://www.uam.es/centros/psicologia/paginas/cpa/paginas/doc/documentacion/rincon/protocolo_ideacion_suicida.pdf: s.n.
- Whittaker, R. et al., 2012. MEMO—A Mobile Phone Depression Prevention Intervention for Adolescents: Development Process and Postprogram Findings on Acceptability From a Randomized Controlled Trial. *J Med Internet Res*, 1(14), p. e13.
- Wirtz, S., Jakobs, E. M. & Ziefle, M., 2009. *Age-specific usability issues of software interfaces*. s.l., s.n.
- Yoshida, A., Hagita, Y., Yamazaki, k. & Yamaguchi, T., 2002. Which do you feel comfortable, interview by a real doctor or by a virtual doctor? A comparative study of responses to inquiries with various psychological intensities, for the development of the Hyper Hospital. *IEEE*, pp. 370-74.
- Zanbaka, C., Goolkasian, P. & Hodges, L., 2006. *Can a virtual cat persuade you? The role of gender and realism in speaker persuasiveness*. Montreal Quebec Canada, s.n.