The final publication is available at

http://doi.org/10.1016/j.ijhcs.2015.11.003

Additional Information

# A Novel Approach to Improve the Planning of Adaptive and Interactive Sessions for the treatment of Major Depression.

Adrián Bresó[1*], Juan Martínez-Miranda[5], Elies Fuster-Garcia[1,2] and Juan Miguel García-Gómez[1,2,3,4]

[1]Grupo de Informática Biomédica, Instituto ITACA, Universitat Politècnica de València, Spain

{adbregua, juanmig, elfusgar}@upv.es

[2]Veratech for Health, S.L., València, Spain

[3]GIBI230 (Grupo de Investigación Biomédica en Imagen), Instituto de Investigación Sanitaria (IIS) hospital la Fe, Spain

[4]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, Bulevar Sur S/N, Valencia 46026, Spain

[5] Centro de Investigación Científica y de Educación Superior de Ensenada, Unidad de Transferencia Tecnológica (CICESE-UT3), Tepic, Mexico

jmiranda@cicese.mx

*Corresponding author

**Abstract**

Human Computer Interaction (HCI) is a research field which aims to improve the relationship between users and interactive computer systems. A main objective of this research area is to make the user experience more pleasant and efficient, minimizing the barrier between the users' cognition of what they want to accomplish and the computer's understanding of the user's tasks, by means of user-friendly, useful and usable designs. A bad HCI design is one of the main reasons behind user rejection of computer-based applications, which in turn produces loss of productivity and economy in industrial environments.

In the eHealth domain, user rejection of computer-based systems is a major barrier to exploiting the maximum benefit from those applications developed to support the treatment of diseases, and in the worst cases a poor design in these systems may cause deterioration in the clinical condition of the patient. Thus,

a high level of personalisation of the system according to users' needs is extremely important, making it easy to use and contributing to the system's efficacy, which in turn facilitates the empowerment of the target users. Ideally, the content offered through the interactive sessions in these applications should be continuously assessed and adapted to the changing condition of the patient. A good HCI design and development can improve the acceptance of these applications and contribute to promoting better adherence levels to the treatment, preventing the patient from further relapses.

In this work, we present a mechanism to provide personalised and adaptive daily interactive sessions focused on the treatment of patients with Major Depression. These sessions are able to automatically adapt the content and length of the sessions to obtain personalised and varied sessions in order to encourage the continuous and long-term use of the system. The tailored adaptation of session content is supported by decision-making processes based on: (i) clinical requirements; (ii) the patient's historical data; and (iii) current responses from the patient. We have evaluated our system through two different methodologies: the first one performing a set of simulations producing different sessions from changing input conditions, in order to assess different levels of adaptability and variability of the session content offered by the system. The second evaluation process involved a set of patients who used the system for 14 to 28 days and answered a questionnaire to provide feedback about the perceived level of adaptability and variability produced by the system. The obtained results in both evaluations indicated good levels of adaptability and variability in the content of the sessions according to the input conditions.

## 1. Introduction

Human Computer Interaction (HCI) is defined by the ACM Special Interest Group on Computer-Human Interaction[1] (ACM SIGCHI) as a discipline concerned with the design, evolution and implementation of interactive computing systems for human use, and with the study of the major phenomena surrounding them *[1]*. A good user interface facilitates effective communication between the user and the software application, but bad HCI design can cause non-acceptance, non-use of the system and user frustration *[2]*.

Karray et al. *[3]* identified a new generation of HCI systems: i) **Intelligent HCI**, which are interfaces that incorporates at least some kind of intelligence in the perception process with the user in order to respond accordingly (e.g. the use of natural language understanding and recognition of body movements); and ii) **Adaptive HCI**, which allows adjustments of its behaviour to each user at any time on the basis of some form of learning, inference, or decision making *[4]*. Tomlinson et al. *[5]* also defined an adaptive interface as the way of the features a user would find desirable and customisable.

In Adaptive HCI, verbal and nonverbal information (such as facial expressions, posture, point of gaze, and the speed/force used when a mouse is moved or clicked, and bioelectrical signals) can be analysed by the system to infer new knowledge about the user and adapt the functionalities of the system to maximise user acceptance, usability and satisfaction level. This adaptation might be related to the (i) **presentation**, i.e. HOW the interaction is conveyed (such as updating screen colours, sounds, etc. *[6]*), or to the (ii) **content**, i.e. WHAT are the actions to be done during the interaction (such as the content of the conversation with the user *[7]*). Additionally, when we are speaking about systems implemented on the Web, there are several authors *[8] [9]* who added a third adaption feature: (iii) the **structure** adaptation or adaptive navigation support which is the mechanism responsible for changing the appearance of visible links.

One of the applications where a good HCI design is important is in the medical domain because these kind of applications could provide great benefits in supporting tasks related to the treatment of patients. But if the design of these applications is faulty or erroneous, it could be harmful and it could have serious consequences for the users such as treatment abandonment, which could lead to a worsening of the patient's health, or cause a relapse [10]. A particularly critical field is the use of HCI techniques in the

---

[1] http://www.sigchi.org/

mental health area where the discontinuation of treatment increases the risk of suicide and death [11]. The design and development of good HCI systems applied to computer-based psychotherapy must be performed taking into account the particular characteristics of the targeted users, such as their cognitive/behavioural capabilities and limitations [12]. The identification and addressing of these special characteristics is particularly important in patients with Major Depression who have associated distorted and negative thinking, which makes them prone to suffering anxiety, frustration and stress when interacting with computer systems [13]. Significant efforts are still required to develop systems that can be widely accepted and that effectively promote the adherence to computer-based psychotherapy.

In this paper we describe an adaptive HCI framework as the core of a Clinical Decision Support System (CDSS), which in turn is one of the components of a Personal Health System (PHS) developed in the context of the Help4Mood European Research Project. The main objective of the project is to support the remote treatment of people who are recovering from a major depressive disorder. The work presented here is concentrated on the generation of tailored sessions (i.e. WHAT their contents are), based on the analysis of user (objective and subjective) inputs and the planning of the daily interactive sessions. The work performed related with HOW to convey the session contents to the patient is out of the scope of this paper but details can be found in [14] [15] [16].

The hypothesis that conducts our work is that "*the dynamic selection and planning of the activities to be included in daily interactive sessions for the treatment of Major Depression based on a user model would generate better adaptive and varied content. Hence, the generation of personalised and varied content can in turn contribute to facilitating the effective use and adherence from users to the system aiming to support the treatment of major depression*".

The algorithm developed for the planning of the content for the daily interactive sessions is based on the knowledge inferred from (i) objective and subjective data collected from the patient; (ii) the historical data that forms a dynamic model of the user, and (iii) a set of requirements pre-defined by the clinicians. In addition to the content of the daily sessions, our proposed framework also produces periodic summary reports with textual and graphical information reflecting the patient's wellbeing evolution in an easy to digest format for both patients and clinicians. These summaries stimulate joint (clinician and patient) reflection about the evolution and improvements achieved by the patient at the different stages of the treatment [17]. All the different modules that form the complete framework of the system have also been

designed to be smoothly extended with new content that can be included in the sessions or easily adapted to other mental health disorders where a continuous monitoring combined with daily sessions could benefit the treatment.

In order to assess how well our proposed framework is able to generate enough levels of adaptive and varied sessions, we evaluated our system using two different approaches: (1) massive simulations representing daily interactions between the user and the system in order to perform a quantitative analysis using statistical methods; and (2) clinical pilots with real patients (N=9) to collect subjective feedback about the perceived levels of variability and adaptability of the content produced by the framework. The rest of the paper is organised as follows: in Section 2 we present the related work. Section 3 describes the design and implementation of the proposed framework. The evaluation methods are described in Section 4, and the obtained results are showed in Section 5. Finally, Section 6 presents some conclusions and future work.

## 2. Related Work

Human-Computer Interaction emerged in the early 1980s, but in the last decade there have been increasing improvements in the field, producing the development of new methods and technologies. Recent achievements in this area have originated new approaches such as adaptive HCI with applications in several areas. Regarding Web platforms, we can find adaptive user interface focused on web searching such as Kinley's study [18], which examined the relationships between users' cognitive styles and their Web searching behaviour. This study may help to provide an adaptive navigation interface that can facilitate efficient retrieval of the relevant search results. In eLearning systems, the adaptive learning interfaces were used to adapt courses, learning material and activities to the learner's individual situation, characteristics and needs [19]. The development of applications for mobile devices is one of the most popular areas in which adaptive HCI is applied. Mobile applications are complex since they need to provide sufficient features to a variety of users in a restricted space where small numbers of components are available. Some authors had proposed their frameworks for mobile applications to make the interfaces automatically adapted to the users. Using data mining (K-means clustering algorithm) Nivethika et al. [20] adapted the application to the experience level of the user based on user historical interactions. Fukazawa et al. [21] proposed and evaluated a method that ranks (using Ranking SVM -Support Vector Machine-) the menu functions of a mobile application based on user operation history. Bae et al. [22] proposed an adaptive transformation

framework that automatically adjusts the content and the appearance to different devices (different sizes and capabilities).

The health care domain is a significant area in which adaptive HCI can be of great benefit to eHealth systems, where  one of the main challenges is the management and delivery of critical information in a way which is easy to understand for heterogeneous groups of patients (different individual abilities, interests, and needs) [9]. One of the main goals in this domain is to enhance the acceptability and usability of health care applications, thereby contributing to a better personalisation of the system outputs for a particular user (i.e. patient, doctor or both).

Adaptive techniques in HCI have been integrated in several Clinical Decision Support Systems (CDSS) in all stages of medical treatment and in most medical areas. Sherimon et al. [23] presented an adaptive questionnaire for diabetic patients based on ontologies, semantic profiles, guidelines, and risk assessment. This questionnaire adapts itself based on the patient's medical history. Bental et al. [24] describe an adaptive medical system for patients with cancer that uses both content and navigation adaptation. The content of the presented information is adapted to the patient's situation and disposition, and the process of illness and treatment. The system proposed by Francisco-Revilla et al. [25] supports adaptive medical information delivery of different medical tasks for users with different levels of expertise. This system supports three tasks: description of medical procedures, supporting the diagnosis, and providing information on health concerns. The work described in Giorgino [26] presents a prototype of a home monitoring system for hypertensive patients through a dialogue on the telephone with an intelligent system. The system implements an automatic speech recognition module in order to collect data about their health status (such as blood pressure, heart rate, or weight) to infer an evaluation using medical guidelines, and performed the content of the conversation. Another related work was presented by Kharat et al. [7] in which they presented a system able to adapt the conversation content with humans based on emotion recognition from facial expressions using neural networks.

Several Knowledge-based techniques in Artificial Intelligence are available to represent the knowledge in a DSS. One of the most popular techniques used for problem solving in intelligent systems is the use of recommendation systems such as Case-Based Reasoning (CBR) which uses past store cases to solve a new problem by recalling similar cases; and Constraint-Based Recommendation which is based on an explicitly defined set of variables and constraints. Another popular technique is the Rule-Based System

(RBS) which solves problems by facts and rules derived from expert knowledge. We can find RBS applied in different health domain, such as in [27], in which a RBS was developed and evaluated in the prevention and treatment of diabetes mellitus. In [28] an adaptive RBS implemented an iterative technique based on previous experience to improve clinical-decision making. So a rule-based system can only be implemented if comprehensive knowledge is available. It is possible to use a hybrid of these systems such as in Ekong et al. [29], where neural networks, fuzzy logic and CBR are combined to model a DSS for the diagnosis of depression disorders. Another hybrid example is provided by Wang et al. [30], where he combines CBR and RBS with fuzzy theory for planning treatments in young people with mental disorders. So, the proper acquisition, representation and management of this knowledge are mechanisms responsible for the behaviour of the system, and which determine the efficiency of the HCI.

In the context of adaptive HCI applied in computer-based psychotherapy for the treatment of anxiety or depressive disorders, we can find systems that implement Cognitive Behavioural Therapy (CBT) as the key component of the session's content such as Beating the Blues (http://www.beatingtheblues.co.uk/) [31], Overcoming depression [32], or MoodGym (https://moodgym.anu.edu.au) [33]. The content of the sessions is predefined on sequencing CBT activities where the user continuously follows the path that the CBT therapy sets and patient responses do not influence the planning of the disclosed activities. One of the initiatives that developed a system similar to the one presented here is the SimCoach project (http://www.simcoach.org/) [34]. SimCoach is a web-based virtual character that aims to help military personnel and their family members in different areas related to mental health including depression, post-traumatic stress disorder, brain injuries, substance abuse and the prevention of suicide amongst others, by offering expert advice and healthcare information. SimCoach does perform dynamic planning of session content before the next activity is disclosed, depending on the patient response. The system provides a text analysis mechanism applied to patient inputs. The system detects keywords associated with different texts / activities and infers the next most suitable activity. A key difference between our own system and SimCoach is that in our framework, the content of the session is not only adapted based on the patient's responses but on three main factors: (i) actual patient responses, (ii) historical data, and (iii) clinical requirements. Moreover, the processing and analysis of patient inputs are performed differently according to the type of data received (i.e. questionnaire scores, signals of physical and sleep activity coming from sensor devices, agreement with suggested activities or the selection of negative thoughts). Finally, Lisetti et al. [35] presented and evaluated an empathic VA that aimed to

increase the effectiveness of behaviour change in patients with excessive alcohol consumption, by means of user engagement and motivation. The main differences between our framework and Lisetti's work is that they analyse only subjective patient input data (such as questionnaires) while in our work we also analyse objective data collected from a set of sensor devices to provide useful and personalised treatment-related activities.

## 3. Design of a Modular Architecture

The design of our system has been performed adopting a user centred design (UCD) methodology by involving a set of patients, clinicians and caregivers with expertise in bringing continuity of care at all levels of healthcare delivery. Figure 1 shows the general architecture of the proposed Personal Health System, which is formed by three separate but interrelated, layers designed to facilitate scalability, flexibility, and maintainability. The layers include:

(i) **The data reception layer** (or input layer), responsible for managing the inputs from the patient, including objective data acquired by a set of actigraphy sensor devices, and subjective data obtained from patient responses to standardised questionnaires.

(ii) **The data processing layer**, which implements the core of the Clinical Decision Support System (CDSS) that analyses the input data and produces the content of the daily sessions. This layer is also responsible for collecting and summarising the information used to construct a weekly summary report.

(iii) **The data transmission layer** (or output layer), responsible for preparing and sending the content of the session to be used by a Virtual Agent to interact with the user. Through this layer the weekly summary report is also sent to the clinician, containing relevant information about the patient's wellbeing evolution.

The work presented in this paper is focused on the Data Processing Layer as the mechanism responsible for producing the adaptive contents of the session according to the patient's detected condition. The importance of this layer from a HCI perspective is how the content that will be communicated to the patient is produced and managed. The content must be sufficiently personalised and adapted to user needs in order to better engage the user. The Data Processing Layer is composed of four modules (their main features are listed in Table 1) that generates the adapted sessions' content:

(a) **The Data Analysis Module (DA)** receives and analyses the signals from the sensor devices to infer clinical findings such as *Low Activity* or *Restless Sleep*; The analysis performed by the DA module includes the following steps:

1. Fusion of the actigraphy signals obtained by multiple sensors using a novel multi-sensor fusion methodology described in [36].

2. Detection of missed data, and segmentation of sleep periods based on a twostep threshold-based strategy described in [37].

3. Generation of daily activity patterns based on FDA formalism [38] as described in [39].

4. Detection of anomalous activity signals based on the K nearest neighbour (Knn) algorithm [40].

5. Generation of enriched comparative plots of daily actigraphy patterns proposed in [39].

All the results of the DA module are stored and are the input used by the rest of the modules.

(b) **The Knowledge Extraction Module (KE)** implements a Rule-Based System that transforms all input information into clinical coded concepts. This module uses the JESS Rule Engine[2] ("Java Expert System Shell") as the inference mechanism. The selection of an RBS was mainly due to the fact that the knowledge representation -from clinical experts- is readily achievable using facts and rules. A large set of simple rules can infer complex behaviour, ensuring good scalability and maintenance.

An example of the KE output is based on the Patient Health Questionnaire (PHQ-9), which is a specific self-report questionnaire to assess depression severity during the past 15 days. Question nine on PHQ-9 is related to self-harm and suicidal thoughts. The *Algorithm* 1 shows an example of a KE rule encoding the representation of patient's response to the PHQ-9 question. The first part of the rule calls a Java function in order to assess if patient suicidal risk exists (using question number 9) or if the overall result of the questionnaire reveals a negative tendency in comparison with previous results. In either of the two cases, the KE infers the "*Deterioration_of_status*" clinical concept in order to alert a negative condition detected in the patient, which will result in the execution of a crisis plan. The second part of the rule assesses the overall result of the PHQ-9 and classifies the level of depression into one of the SNOMED-CT based clinical concepts: Mild, Moderate, or Severe.

---

```
;;------------------------------------------------------------------------
;; RULE KE_PHQ9_QuestionNumber_9_Received
;;------------------------------------------------------------------------
(defrule KE_PHQ9_QuestionNumber_9
  ?l <- (Data{name == "Request_PHQ-9_9"}(idSession ?l_IdSession) (value ?value)(value2 ?value2))
  ?ll <- ( variable (name "KE_PHQ9_Summatory") (value ?valueSum))
  (test (eq ?l_IdSession(Utils.getIdSession)))
  =>
(if (Utils.isEmergencyPlanEnabled (?value ?*THRESHOLD_VALUE_PHQ9_QUESTION*) then
        (bind ?d(new Data))
        (?d1 setName "Deterioration_of_status")
        (?d1 setCode "390772001")
        (Utils.setDate ?d1)
        (Utils.setIdSession ?d1)
        (?d1 setOntology "SnomedCT")
        (add ?d1)
)
;;Assessment of the questionnaire result
(bind ?d(new Data))
(?d setName "KE_PHQ9_Completed")
(if (and(>= (+ ?valueSum ?value) 0)(<= (+ ?valueSum ?value) 5)) then
        (?d setValue2 "MILD DEPRESSION")
        (?d setCode "310495003")
)
(if (and(> (+ ?valueSum ?value) 5) (<= (+ ?valueSum ?value) 10)) then
        (?d setValue2 "MODERATE DEPRESSION")
        (?d setCode "310496002")
)
(if (and(> (+ ?valueSum ?value) 10) ) then
        (?d setValue2 "SEVERE DEPRESSION")
        (?d setCode "310497006")
)
(Utils.setDate ?d)
(Utils.setIdSession ?d)
(?d setOntology "Snomed")
(add ?d)
(retract ?ll)
(retract ?l)
)
```

*Algorithm 1: Example of coded JESS rule in the Knowledge Extraction Module. This example is used to assess whether a suicidal risk or a negative tendency in the patient's depression level exists, inferring the "Deterioration_of_status" concept coded in SNOMED CT as 390772001. Additionally, this rule assesses the depression level and creates a PHQ-9 result concept coded in SNOMED CT ("Mild": 310495003, "Moderate": 310496002, and "Severe": 310497006) based on the current user response.*

(c) **The Knowledge Inference Module (KI)** is a similar RBS to the one implemented in the KE module but with a different purpose. The KI module infers a set of recommended tasks to be suggested by the system to the patient during the course of the session. These tasks represent specific dialogue acts used by the Virtual Agent during the interaction with the patient (see [16] [14]). For example, if the KI receives the "*Restless Sleep*" clinical concept from the KE,the KI will infer the task "*Provide_Sleep_Info*" used to provide information about recommendations to be followed that would help the patient to get better quality sleep. Other information used by the KI includes clinical specifications related to the treatment coded in

the rules of the KI. For instance, one of the tasks offered to the patient is the setting of at most three activities to carry out during the current week (e.g. "go shopping", "call a friend", "take a walk", etc.). When the system detects that the patient has configured a weekly activity plan ("*Configure_Activity_Plan*"), the KI will provide the patient with some reminders about the selected activities ("*Ask_For_Planned_Activities*") and obtain information about whether the patient has performed the activity plan and how many times during the week those activities were carried out.

Finally, the historical data retrieved from the user model is also taken into account to infer new tasks. For example, the system can launch reminders such as "*Activity_Reminder*" or "*Welcome_Reminder*", which are used to suggest the daily use of the system and execution of activities to the user when the system detects that the last data received from the actigraphy sensor or the last user login into the system was more than 3 days ago.

(d) **The Session Planner Module (SP)** manages the set of tasks inferred by the KI. The SP is the mechanism responsible for planning the patient's activities during the daily sessions. The flexibility and dynamism of the SP adapt the content of the daily session at each interaction cycle depending on the current responses from the patient. It is able to make a new plan if necessary by adding, removing and/or changing the order of the activities in real time, avoiding the repetition of sessions with exactly the same content and thus avoiding routine work by the patient.

The SP plans the most adequate daily session for the patient based on:

1.  The set of inferred tasks by the KI.

2.  The user model, which contains information about previous and current sessions (such as the periodicity of the executed activities, banned activities, or the current selection on session length).

3.  The pre-defined settings that clinicians configure according to the most appropriate treatment for each patient (or set of patients). Clinicians can define an adequate scenario, establishing -for each task- the number of times (minimum and maximum) that the task needs to be executed, its priority, and constraints or dependencies between tasks. Some tasks have special characteristics, such as the *interruption tasks* (e.g. the execution of a Crisis Plan triggered by a request from the user in the GUI or by a low score detected in PHQ-9). If the SP detects one of these tasks, the current plan is deleted and the interruption task is executed immediately.

## 4. Evaluation Method

Our research method and functional evaluation of the developed framework has been motivated by answering the following addressed question: "*Is our developed architecture able to generate a sufficient degree of adaptive and varied sessions that can contribute to increasing and maintaining the interest of the users?*". For this purpose, we conducted two different approaches: the generation of simulated data, and the involvement of patients to collect their feedback about the system. In the simulation process, a range of synthetic data was produced to perform a deep statistical analysis of different possible outputs produced by the data processing layer of the system. In the evaluation where a set of patients was involved, our proposed framework was integrated into a full system used in two clinical pilots.

### 4.1 Evaluation Based on Simulations

### 4.1.1 Definition of Scenarios

The methodology to assess the levels of adaptability and variability in the content of the sessions produced by our proposed framework has included the definition of two scenarios based on clinical requirements: **restrictive** and **flexible** scenarios. The difference between the two scenarios is based on how some clinical requirements (associated with the support of the treatment) are pre-defined for the generation and planning of the different activities that form a session. Both scenarios are based on the requirements defined by the clinicians of the Help4Mood project. The settings of the flexible scenario are less stringent than the restrictive scenario, i.e. the number of times that an activity needs to be offered to the patient or how these activities should be ordered during a session is more open than in a restrictive scenario (Table 2 lists the settings used during the generation and planning of the actions –dialogue acts– for both scenarios).

The final planning and selection of the specific content to be included in a session is dependent on these pre-defined settings and the best way to observe the differences produced from the dynamic input data is the comparison of the sessions produced in the two different scenarios. This comparison will help to assess the level of adaptability in our framework to the constraints defined by the clinicians. The rationale behind the definition of different scenarios is the generation of different content of the sessions according to the particular preferences of each clinician. For example, while some specialists would prefer the inclusion of a large number of different activities to support the treatment, some others would prefer the execution of a small number of more focused, activities based on the specific condition of each patient. Bearing this in mind, for the functional evaluation of the framework we have set the restrictive scenario

defined by a high number of constraints in the relative order and dependencies between tasks, as well as in the periodicity and priority of the tasks. The flexible scenario was defined with a minor number of constraints. It is expected that the restrictive scenario would provide a low level of variability in the content of the sessions compared to the flexible scenario due to the differences in the pre-defined settings. An extremely restrictive scenario could generate static daily sessions (i.e. producing exactly the same content in every session). On the other hand, an extremely flexible scenario could generate almost fully random sessions based only by the user inputs generated during the interaction.

### 4.1.2 Variables and Evaluation Space for Simulations

One of the goals in the evaluation process was to demonstrate that our framework is able to produce different levels of adaptive and varied content according to the settings in the different scenarios A set of simulations were executed representing the interaction between the system and the user through the generation of random values in the simulated responses from a patient. The random values used as patient responses were based on the range of values that a patient could give in a real interaction with the system. For example, if the virtual agent of the system asks "How is your mood today?" –which correspond to the Daily Mood Check 1 activity-, the simulated response generates a random value between 0 and 100, the same scale that a patient needs to choose during a real interaction. In other cases, we used categorical responses, such as when the virtual agent asks: "Which session length do you prefer?", corresponding to the activity *Select Type Session*. In this case a random value is generated representing the available "Long", "Medium", or "Short" responses. In this way we have generated a set of simulations containing values that a patient could select during a real interaction. The evaluation space corresponds to the multivariate combination of answers to all questions that might make sense given the context of the patient.

During the execution of the simulations we used 19 tasks (see *Table* 2) and 31 subtasks. A task could be formed by one or more sub-tasks –i.e. different dialogue acts used to provide the patient with all the information to execute the task. For example the *Introduce Relaxation Exercise* is formed by three subtasks: *Select Voice* (the patient selects a pre-recorded voice which provides the instructions of the relaxation exercise), *Preparing Relaxation Exercise* (the patient receives the information required to start the exercise), and *Perform Relaxation Exercise* (the dialogues that represent the execution of the exercise). In order to obtain smoothed distributions of our results, we have executed a total of 20.000 simulations of interactive sessions (10.000 of them with the restrictive scenario and 10.000 with the flexible scenario).

After carrying out simulations, we conducted the assessment of the adaptability and variability levels of the content produced in the session. For our evaluation purposes, we define the adaptability as how much the content of a session (activities/suggestions/questionnaires offered to the patient) can change dynamically during the interaction according to the current and past information received and inferred about the patient's condition. The variability was defined as how the order of the content is offered depending on the user's actions during the interaction and the set of restrictions defined by the clinicians.

## 4.2 Evaluation with Real Patients

### 4.2.1 Definition of Pilots

In order to complete the assessment, an evaluation with real users was performed. Two incremental pilots were defined by the clinical staff of the Help4Mood project according to the ethical and clinical requirements of Help4Mood project. The main aim of these pilots was to test the feasibility of deploying H4M in different clinical contexts. The results presented in this work are only those focused on the subjective feedback collected from the participants about the perceived variability and adaptability of the session's content produced by our proposed frameworks.

In the first pilot seven real depressed patients from Romania and Spain were enrolled. For two weeks, all of them used the full system at their homes. In pilot 2, two patients from Spain took part and were asked to use the system on daily basis for four weeks (see Table 3). The inclusion and exclusion criteria for enrolling in the pilots were:

- Participants were aged between 18 and 64 inclusive.
- Major Depressive Disorder (MDD) as primary diagnosis.
- Absence of other mental disorders (bipolar, psychotic, or panic).
- Beck Depression Inventory II (BDI-II) score above 9 and below 31.
- Participants were recruited on a voluntary basis.
- Participants lived at home, in the community.
- Patients were free from pharmacological and therapeutically treatments.

In both pilots the clinical experts set the framework with a restrictive scenario. The clinicians established the majority of the tasks, as prescribed and strictly. This fact, together with the low number of available tasks in the pilot 1 (N=17), is expected to cause a low variability and adaptability of the system. Nonetheless, in the second pilot two new tasks are added so the variability and adaptability should show a small improvement. These new two additions were the relaxation exercises "*Introduce Relaxation Task*"

(ID=12), and the setting of a plan of activities to perform during a week "*Ask For Planned Activities*" (ID=14).

**4.2.2 Evaluation Procedure**

At the end of the pilots, every participant filled in an eleven-item questionnaire to collect the perceived usefulness of different functionalities of the system. Regarding the perceived level of variability and adaptability, the questionnaire included two questions (see Table 4): Q1 for the assessment of the adaptability (length of the sessions) and Q2 for the assessment of the variability in the content offered by the system. We used the 3-point Likert scale (+1: adequate; 0: neither; and -1: non-adequate) in order to compare the results of the two pilots.

**5. Results**

**5.1 Results from Simulations**

**5.1.1 Adaptability**

Taking into account that a key characteristic of people with depression is the loss of interest in doing things, and even simple daily life activities would represent a major effort, the most relevant parameter to assess if the level of adaptability of the system is adequate or not according to the patient condition is the length of the sessions. This allows us to know if the length of the sessions needs to be adapted according to the patient's stamina. Hence, we evaluate the adaptability level as the number of planned tasks during a session. The implementation of our framework into the Help4Mood system to learn the patient's stamina is through a direct question about what type of session's length the patient would prefer (*Select Session Type* task). According to the patient preference our framework adapts the content, and thus the length of the session, according to the scenario and to the model of the user. In the flexible scenario it is expected that the length of the session will be more conditioned by the user's answer, because only functional restrictions and minimum clinical restrictions have been set in the system. However in the restrictive scenario it is expected that the length of the session will not only be conditioned by the user's answer but also by the restrictions set in the system based on therapeutic and clinical indications. It is expected that *long* sessions will be longer in the flexible than in the restrictive scenario, and that *short* sessions will be shorter in the flexible than in the restrictive scenario, mainly due to the different settings

related to the values associated with the minimum and maximum (*minExecutions* and *maxExecutions*) number of executions of the activities per week in the two scenarios.

The analysis of the adaptability level is based on the differences between the probability distributions of the short, medium and long sessions obtained from the simulations in both scenarios. These distributions have been obtained using the Kernel smoothing function estimate [41] based on a normal kernel smoother. From the obtained distribution of the total amount of planned activities we can identify in the graphs of Figure 2 the three different types of sessions.

As we observe in the plots, there is a difference in the length of the session depending on the defined scenario. These differences originate from the settings associated with the periodicity of the activities defined in the *maxExecutions and minExecutions* parameters set with different values in each scenario. In the short and medium length sessions, the number of activities planned in the restrictive scenario produce slightly longer sessions than in the flexible scenario (see the slope of the probability distribution functions in the first two plots of Figure 2). The reason for this difference is that in the restrictive scenario the values of the parameter *minExecutions* are greater than in the flexible scenario (representing a strong requirement that specific activities must be performed a minimum number of times during a week). Therefore, at some point the session length is extended to meet the requirement to perform the specified activities the minimum number of times. On the other hand, *long* sessions trend to be shorter in the restrictive scenario than in the flexible scenario. The reason is that the Session Planner module -regardless of the scenario- always includes all the tasks inferred by the KI, except those tasks that have been executed the number of times defined in the *maxExecutions* parameter or that fail to fulfil any other criteria. In the restrictive scenario the values of the *maxExecutions* parameter are smaller than in the flexible scenario (representing a strong condition that some activities should be executed only a few times during a week). Therefore, for *long* sessions the number of activities planned in restrictive scenarios is smaller than in flexible scenarios.

The differences in session length between the two scenarios are also reflected in terms of the variation in the mean values of the distributions as presented in the following Table 5.

To assess the differences between the two scenarios for each of session types we have used the nonparametric Mann-Whitney U test. For the three session types (long, medium and short), the differences

16

in session duration between flexible and restrictive scenarios have been found significantly different, obtaining p-values under 0.01.

The obtained results support the expected results: in restrictive scenarios the content of the session is highly influenced by the clinical and therapeutic restrictions or preferences, whilst in flexible scenarios the length of the session is more influenced by the actions and inputs received from the user than by the pre-defined settings.

### 5.1.2 Variability

One of the key aspects in the provision of daily sessions to patients is the capability of the system to offer different content that minimises the risk of system use discontinuation due to the repetitive and routine execution of exactly the same sequences of activities. One strategy to provide varied content during the sessions, even when the offered tasks could be the same in order to meet the pre-defined clinical constraints, is to offer the activities in a different order during the interaction with the user. This way, the patient will address the offered tasks at different moments during the session and the feedback obtained from the Virtual Agent will also change depending on the results of the activities received from the patient. In the simulations, we have also analysed the level of variability produced by our proposed framework.

As in the evaluation of the adaptability, we expected that the level of variability would depend heavily on the configuration that the clinician had defined. Since the framework has been designed to be configured either with a high level of flexibility or moderately static, we have used the same two scenarios again to assess the level of variability that the framework can produce. Comparing the different positions in which the same task is planned for each of the scenarios, we can see some differences. In general, there is more variability in the task position in a flexible scenario than in a restrictive scenario. In a flexible scenario the probability of transition from one activity to another is quite similar for all the activities, except those transitions that are subjected to functional restrictions. The inclusion of clinical and therapeutic preferences in the settings of the system reduces the number of choices or degrees of freedom in the system, potentiating some transitions above others. This behaviour means that the level of variability in a flexible scenario will decrease with the inclusion of the clinical preferences or restrictions.

The Figure 3 shows two heat maps, one for each scenario, where the numbers of transitions from one activity to other during the simulations are represented.

We can see that in both scenarios the transition from tasks 1 and 2 to task 3 occurs with high frequency. The task 1 and task 2 represent the dialogue acts to welcome the user at the beginning of each session while task 3 is the first question of the daily mood check questionnaire. Since the clinicians defined that the daily mood check questionnaire must be the very first task (immediately after the welcome) in each session, we can see in the Figure 3 almost identical task transitions in the two scenarios (indices from 1 to 9). In the two scenarios there is a low variability in the order of the tasks corresponding to the four questions belonging to the *daily mood check questionnaire* activity.

Beyond these initial activities, we can start to see the differences between the flexible and the restrictive scenarios on the top right area of the graph. In the flexible scenario, the transitions are more homogeneous and there are not many defined patterns and we can argue that in a flexible scenario there is more variability between sessions than in restrictive scenarios. For the analysis, we have focused on the transitions between indices from 9 to 18, which corresponds to the variable transitions, excluding the transitions that are forced by functional requirements.

As can be seen in Figure 4, in the flexible scenario most of transitions have a similar probability while in the restrictive scenario there is a clear occurrence of banned and recurrent transitions. When analysing the differences in number of low probability transitions (i.e. transitions with less than 2% probability of occurring), we can see that in the flexible scenario only 30% of the transitions were low probability transitions, while in the restrictive scenario the number of low probability transitions increases to 64%. These experiments show how the variability of the session significantly decreases when adding a significant number of restrictions in the settings of the system.

We have also analysed the number of repeated sessions obtained during the simulations for the two scenarios. Repeated sessions are those that contain exactly the same number and in the same order as the activities produced in the session planner. From this analysis, we can see that in the flexible scenario the level of variability is greater than in the restrictive scenario (see *Table* 6). The results obtained in the restrictive scenario show 11.87% of repeated sessions. We can also see that in short sessions, the variability level is lower than in medium and long sessions, which is justified for the smaller number of planned activities.

**5.2 Results from Pilots**

**5.2.1 Adaptability and Variability**

The results obtained in pilot 1 showed that the patients considered that the system had a low level of variability and adaptability. These results were not unexpected due to the low number of tasks used during the sessions and the configured restrictive scenario. The results obtained in the second pilot outperformed the results obtained in the first pilots in both the perceived level of variability as well as adaptability. This improvement in the perception of the participants was mainly due to the addition of the two new tasks that produced more combinations of different content and length of the sessions each day. These results (see Table 7) demonstrated that independent of the configuration of the scenario, the more available tasks, the better the levels of variability and adaptability produced by the framework, as also demonstrated with the synthetic data produced in the simulations presented in Section 4.1.

## 6. Discussion and Conclusions

In this work, we present an adaptive and flexible framework to produce the content of interactive daily sessions applied for the treatment of Major Depression. This framework has been designed with the aim of improving the acceptance of a system to support the remote treatment of major depression in the targeted users, which in turn would help to increase the adherence to the treatment.

In the design and implementation of HCI addressed to support the self-treatment of patients, there are several recommendations to follow in order to avoid frustration and loss of interest. These recommendations include the promotion of the variability and the adaptability of the system outputs to the patient's condition. The proposed framework provides a mechanism that produces a personalised adaptation and offers a high level of variability in the content of the sessions based on a continuous and dynamic planning of the activities to offer to the patient. The planning adapts the content of each session according to the patient's condition and input actions at each interaction cycle. The session can vary the content, the length and the order of the tasks to produce different sessions every day. The management of the produced sessions is based on the patient's direct answers, the historical data, and the preferences of the clinicians encouraging better patient adherence to the treatment.

The functionality of the framework has been evaluated to observe the produced levels of variability and adaptability in the daily sessions. The evaluation has been performed using simulations to represent patients' inputs and the corresponding system responses. After analysing the results, and considering that a session is composed of an average of 18 tasks (13 for short session, 17.5 for medium sessions, and 24 for long sessions), we can conclude that there is a 30% of overall adaptability between planned sessions in our

framework. Regarding variability, we obtained only 10.92% of repeated sessions. We can argue that the presented framework provides good levels of variability and adaptability according to the inputs received from the patient.

Due to the high dependency between the planning process and the configuration of the system set by the clinicians, we performed simulations using two different configurations (which we called scenarios): a restrictive scenario in which we set a high number or restrictions; and a flexible scenario in which we set a more open configuration (see *Table* 2). We have shown that when the system is configured in a more restrictive way, the level of variability and adaptability is reduced. In contrast, when a more flexible configuration is set in the system, the level of variability and adaptability is improved.

The adoption of different configuration scenarios enables the system to better adapt to the needs of the clinicians and their patients. We can ensure that our framework provides a *sufficient degree of adaptive and varied sessions, allowing the personalisation of the interactive sessions in order to improve the user experience.* The proposed framework can be used to support computer-based psychotherapeutic interventions in patients who require high restrictions in the generation of different session contents guided by clinical settings, or in patients who need a more flexible treatment. The clinician is responsible for setting up the tasks and conditions to generate the more adequate content for the treatment, including the frequency and restrictions of the different activities based on the patient condition or the protocols of the clinical institution.

Additionally, a complementary evaluation was performed, in which we collected the feedback of a set of patients who used the system on a daily basis for two to four weeks. The collected feedback confirms that even when the system uses a restrictive scenario, if the available set of tasks to offer the user is sufficiently extended, the variability and the adaptability of the system is improved. Finally, we can conclude that our framework depends heavily on the configuration of the scenarios and on the set of tasks that it can plan.

For future work we are considering different actions to improve the current framework. First, we will extend the evaluation with more participants in a clinical setting, in order to collect a greater set of samples about the content and length of the sessions to be able to better assess the overall acceptability and usability of the proposed framework.

**References**

[1] C. Rusu, V. Rusu and S. Roncagliolo, "Usability practice: the appealing way to HCI.," in *In Advances in Computer-Human Interaction, First International Conference on (pp. 265-270). IEEE.*, 2008.

[2] K. Bessière, J. Newhagen, J. Robinson and B. Shneiderman, "A model for computer frustration: the role of instrumental and dispositional factors on incident, session, and post-session frustration and mood," *Computers in Human Behavior,* vol. 22, pp. 941-961, 2006.

[3] F. Karray, M. Alemzadeh, J. A. Saleh and M. N. Arab, "Human-Computer Interaction: Overview on State of the Art," *International Journal on Smart Sensing and Intelligent Systems,* vol. 1, no. 1, pp. 137-159, 2008.

[4] A. Sears and J. A. Jacko, Human-computer Interaction: Design Issues, Solutions, and Applications., CRC Press., 2009.

[5] G. Tomlinson and E. Baumer, "Dreaming of Adaptive Interface Agents.," in *ACM Conference on Human Factors in Computing Systems*, San Jose, CA, USA ACM: 2007 - 2012.Amsterdam, 21B, 247-250, 2007.

[6] K. Börner, "Concept-Based, Adaptive Human-Computer Interaction," in *Proceedings of the 9th Irish Conference on Artificial Intelligence and Cognitive Science*, 1994.

[7] G. U. Kharat and S. V. Dudul, "Emotion recognition from facial expression using neural networks.," *In Human-Computer Systems Interaction,* vol. Springer Berlin Heidelberg, pp. 207-219, 2009.

[8] A. Bunt, G. Carenini and C. Conati, "Adaptive content presentation for the web.," *In The adaptive web,* vol. Springer Berlin Heidelberg, pp. 409-432, 2007.

[9] E. Vasilyeva, M. Pechenizkiy and S. Puuronen, "Towards the framework of adaptive user interfaces for eHealth.," in *In Computer-Based Medical Systems, 2005. Proceedings. 18th IEEE Symposium on (pp. 139-144). IEEE.*, 2005.

[10] D. Novick, J. M. Haro, D. Suarez, V. Perez, R. W. Dittmann and P. M. Haddad, "Predictors and clinical consequences of non-adherence with antipsychotic medication in the outpatient treatment of schizophrenia.," *Psychiatry research,* vol. 2, no. 176, pp. 109-113, 2010.

[11] A. Ward, K. Ishak, I. Proskorovsky and J. Caro, "Compliance with refilling prescriptions for atypical antipsychotic agents and its association with the risks for hospitalization, suicide, and death in patients with schizophrenia in Quebec and Saskatchewan: a retrospective database study," *Clinical therapeutics,* vol. 28, no. 11, pp. 1912-1921, 2006.

[12] V. L. Patel and A. W. Kushniruk, "Interface design for health care environments: the role of cognitive science.," *Interfaces,* vol. 4, no. 5, 1998.

[13] S. M. Safford and J. E. Worthington, "Computer anxiety in individuals with serious mental illness.," *Computers in human behavior,* vol. 15, no. 6, pp. 735-745, 1999.

[14] J. Martínez-Miranda, A. Bresó and J. M. García-Gómez, "Modelling Two Emotion Regulation Strategies as Key Features of Therapeutic Empathy," *To appear in SEM2011 / EEA2012,* 2014.

[15] J. Martínez-Miranda, A. Bresó and G.-G. J.M, "The Construction of a Cognitive-Emotional Module for the Help4Mood's Virtual Agent," in *Proceedings of the 1st Workshop on ICT applied to Mental Health*, Spain (1) 2012:34-39, 2012.

[16] J. Martínez-Miranda, A. Bresó and J. García-Gómez, "Look on the Bridge: A Model of Cognitive Change in Virtual Agents," *In Proceedings of the 14th International Conference on Intelligent Virtual Agents (IVA) Springer LNCS,* vol. 8637 (in press), 2014.

[17] M. K. Wolters, C. D. Burton, C. Matheson, A. Bresó, A. Szentagotai, J. Martínez-Miranda, J. Rosell, E. Fuster, C. Pagliari and B. McKinstry, "Help4Mood - Supporting Joint Sense Making in the Treatment of Major Depressive Disorder," in *In Workshop on Interactive Systems in Healthcare (WISH)*, Washington DC, 2013.

[18] K. Kinley, D. W. Tjondronegoro, H. L. Partridge and S. L. Edwards, "Modeling users' web search behavior and their cognitive styles," *Journal of the American Society for Information Science and Technology,* 2014.

[19] S. Graf, "Adaptive Technologies," in *Handbook of Research on Educational Communications and Technology*, New York, Springer , 2014, pp. 771-779.

[20] M. Nivethika, I. Vithiya, S. Anntharshika and S. Deegalla, "Personalized and adaptive user interface framework for mobile application," in *In Advances in Computing, Communications and Informatics (ICACCI), 2013 International Conference on (pp. 1913-1918). IEEE.*, 2013.

[21] Y. Fukazawa, M. Hara, M. Onogi and H. Ueno, "Automatic mobile menu customization based on user operation history," in *11th International Conference on HCI with Mobile Devices and Services.*, 2009.

[22] Y. Bae, B. Oh and J. Park, "Adaptive Transformation for a Scalable User Interface Framework Supporting Multi-screen Services," *Ubiquitous Information Technologies and Applications,* vol. Springer Berlin Heidelberg, pp. 425-432, 2014.

[23] P. C. Sherimon, P. V. Vinu, R. Krishnan, Y. Takroni, Y. AlKaabi and Y. AlFars, "Adaptive Questionnaire Ontology in Gathering Patient Medical History in Diabetes Domain," in *In Proceedings of the First International Conference on Advanced Data and Information Engineering (DaEng-2013) pp. 453-460*, Singapore, 2014.

[24] D. Bental, A. Cawsey, R. Jones and J. Pearson, "Adapting Web-Based Information to the Needs of Patients with Cancer," in *In Proc. of Int. Conf. on AH and Adaptive Web-based Systems, Springer (pp. 27-37)*, 2000.

[25] L. Francisco-Revilla and F. Shipman, "Adaptive Medical Information Delivery: Combining User, Task and Situation Models," in *in SIGWEB Newsletter, Vol 8, Number 1, 1999, ACM Press (pp. 18-20)*, New York, USA.

[26] T. Giorgino, "Automated Spoken Dialog System for Hypertensive Patient Home Management," *International Journal of Medical Informatics ,* vol. 74, no. 2-4, pp. 159-167, 2005.

[27] A. Bresó, C. Sáez, J. Vicente, F. Larrinaga, M. Robles and J. M. García-Gómez, "Knowledge-based personal health system to empower outpatients of diabetes mellitus by means of p4 medicine," in *Data Mining in Clinical Medicine*, New York , Springer , 2015, pp. 237-257.

[28] M. Serhani, A. Benharref and A. Nujum, "An Adaptive Expert System for Automated Advices Generation-Based Semi-continuous M-Health Monitoring," *Brain informatics and health,* pp. 388-399, 2014.

[29] V. E. Ekong, U. G. Inyang and E. A. Onibere, "Intelligent Decision Support System for Depression Diagnosis Based on Neuro-fuzzy-CBR Hybrid.," *Modern Applied Science,* vol. 6, no. 7, p. 79, 2012.

[30] W. M. Wang, C. F. Cheung, W. B. Lee and S. K. Kwok, "Knowledge-based treatment planning for adolescent early intervention of mental healthcare: a hybrid case-based reasoning approach," *Expert Systems,* vol. 24, no. 4, pp. 232-251, 2007.

[31] S. Van Den Berg, D. Shapiro, D. Bickerstaffe and K. Cavanagh, "Computerized cognitive-behaviour therapy for anxiety and depression: a practical solution to the shortage of trained therapists.," *Journal of Psychiatric and Mental Health Nursing,* vol. 11, no. 5, p. 508, 2004.

[32] K. Colby and C. P.M, "Overcoming depression.," *Malibu, CA. Malibu Intelligence Artificial Works,* 1990.

[33] H. Christensen, K. Griffiths and A. Korten, "Web-based Cognitive Behavior Therapy: Analysis of Site Usage and Changes in Depression and Anxiety Scores," *Journal of Medical Internet Research,* vol. 4, no. 1, 2002.

[34] A. Rizzo and B. J. Joseph, "Virtual Reality, A New Therapy Part Two: The SimCoach," *Psychology Today,* 2011.

[35] C. Lisetti, R. Amini, U. Yasavur and N. Rishe, "I Can Help You Change! An Empathic Virtual Agent Delivers Behavior Change Health Interventions," *ACM Transactions on Management Information Systems,* vol. 4, no. 4, p. 19:1, 2013.

[36] E. Fuster-Garcia, A. Bresó, J. Martínez-Miranda, J. Rosell, C. Matheson and J. García-Gómez, "Fusing actigraphy signals for outpatient monitoring," *Information Fusion,* 2014.

[37] E. Fuster-García, J. Juan-Albarracín, A. Bresó and J. García-Gómez, "Monitoring changes in daily actigraphy patterns of free-living patients," in *International Work-Conference on Bioinformatics and Biomedical Engineering (IWBBIO)*, Granada (Spain), 2013.

[38] J. O. Ramsay and B. W. Silverman, Functional data analysis., New York: Springer, 2005.

[39] E. Fuster-Garcia, A. Bresó, J. M. Miranda and J. M. García-Gómez, "Actigraphy Pattern Analysis for Outpatient Monitoring," in *Data Mining in Clinical Medicine* , New York, Springer , 2015, pp. 3-17.

[40] V. Chandola, A. Banerjee and V. Kumar, "Anomaly detection: A survey," *ACM Comput. Surv.,* vol. 41, pp. 3-61, 2009.

[41] A. Bowman and A. Azzalini. , "Applied Smoothing Techniques for Data Analysis," *Oxford University Press Inc.,* 197. New York.

[42] M. K. Wolters, J. Martínez-Miranda, S. Estevez, H. Hastie and C. Matheson, "Managing Data in Help4Mood," in *In EAI Endorsed Transactions on Ambient Systems*, Vol. 13 (01-06): e2.
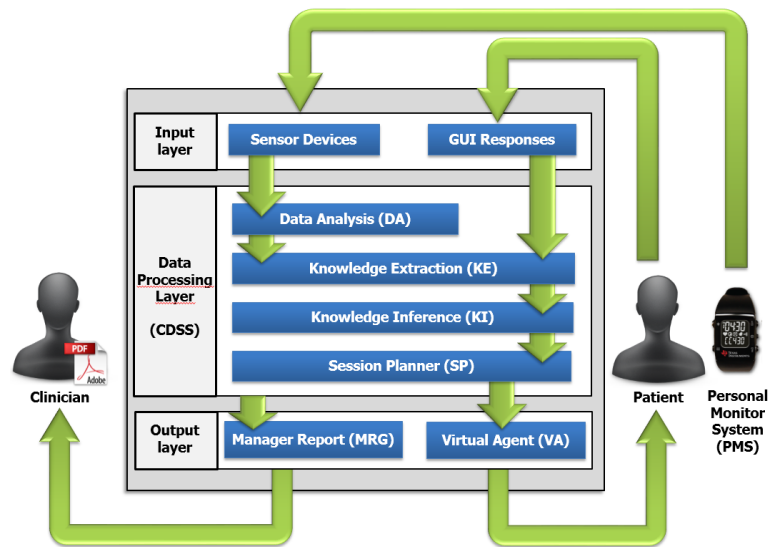
*Figure 1: General architecture of the proposed Personal Health System to provide support in the treatment of severe depression. The patient interacts with the system by means of a Virtual Agent. The activity sensor devices collect and send the acquired data to the system. The content of the daily sessions is generated and continuously adapted in the data processing layer, which implements a Clinical Decision Support System (CDSS). Finally this content is transmitted to the patient and clinician through the output layer.*
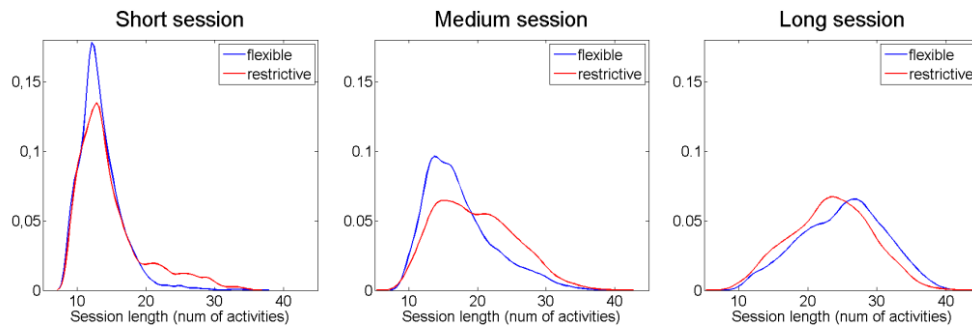
*Figure 2 Probability distributions of the duration of short, medium and long sessions obtained from the simulations in the two scenarios (the blue line shows the flexible scenario, the red line shows the restrictive scenario).*
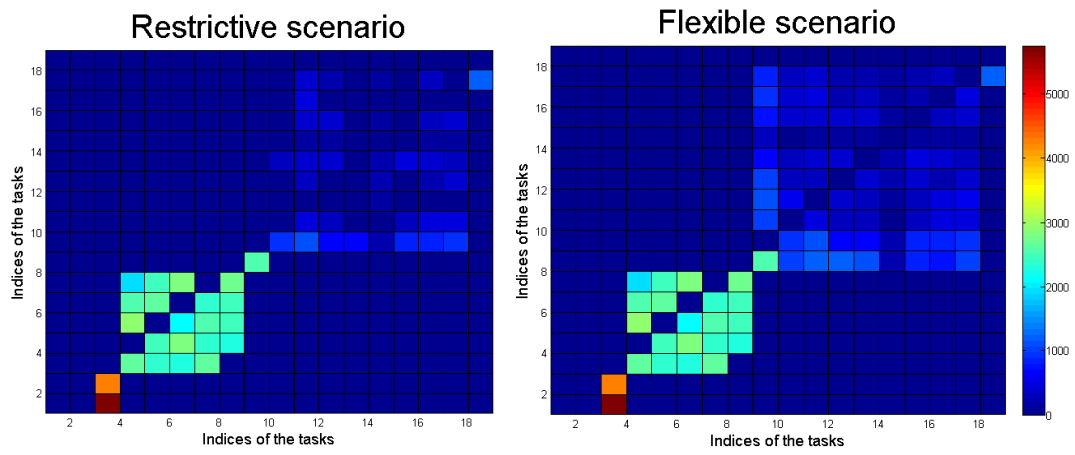
*Figure 3 Heat maps with the number of transitions made from one activity to another during the simulations. The heat map on the left is obtained from the flexible scenario, while the heat map on the right is obtained from the restrictive scenario. The indices of the axis are labelled in Table 2.*
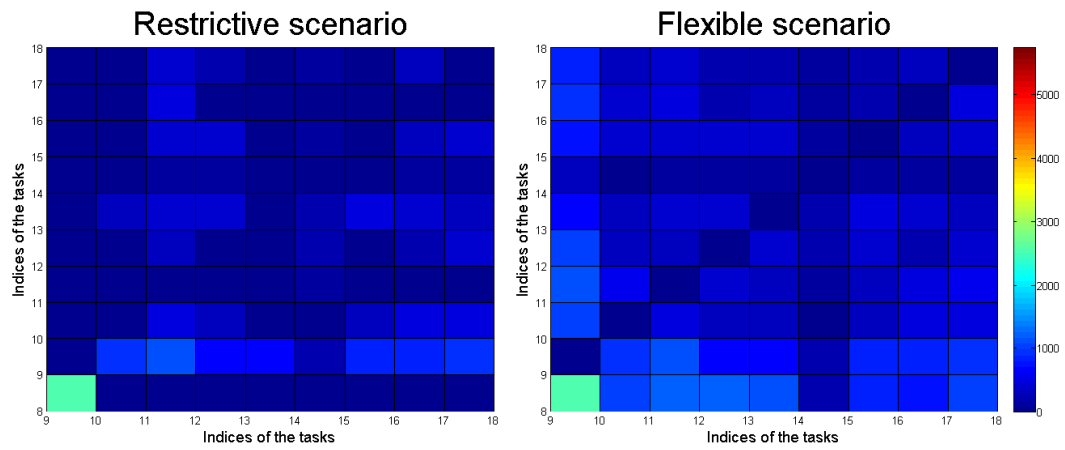
*Figure 4. Detailed view of the region in the heat maps that shows the variable transitions between activities. The detailed heat map on the left is obtained from the restrictive scenario, while the detailed heat map on the right is obtained from the flexible scenario. The indices of the axis are labelled in Table 2.*

| Module | Main Task | Input | Output |
|---|---|---|---|
| Data Analysis (DA) | • Fusion of the actigraphy signals obtained by multiple sensors [36] <br> • Detection of missed data, and segmentation of sleep periods [37] <br> • Generation of daily activity patterns [39] <br> • Detection of anomalous activity signals [40] <br> • Visualization [39] | • Patient's activity data coming from different sensor devices (wrist-watch, phone, under-mattress, or key ring) | • Daily activity models (physical activity and sleep patterns) for specific groups of days (such as working days, weekends, or days in which the patient is under specific stages of the treatment) <br> • Detection of possible crises or relevant events in the future <br> • Set of graphical plots and statistical calculations to be included in the periodic summary report |
| Knowledge Extraction (KE) | • Infer clinical concepts <br> • Code clinical concepts | • DA findings <br> • Patient responses <br> • User Model (demographical + current and historical data) <br> • KE clinical knowledge coded in if-then rules <br> • KE knowledge base | • Clinical concepts which are coded using an internal format based on the SNOMED-CT terminology (such as Mild Depression: 310495003, or Restless Sleep: 12262002) to facilitate the interoperability among the rest of the system's components and with external data repositories [41] |
| Knowledge Inference (KI) | • Infer a set of recommended tasks | • KE findings (Clinical concepts) <br> • Patient responses <br> • User Model (demographical + current and historical data) <br> • KI clinical knowledge coded in if-then rules <br> • KI knowledge base | • Set of tasks, which include the suggestion of specific activities, some reminders, the administering of questionnaires to collect more data from the patient or exercises based on Cognitive Behaviour Therapy (CBT) to help the patient in the identification and reflection on thoughts and experiences |
| Session Planner (SP) | • Interruptions detector <br> • Task Classifier <br> • Task Clustering <br> • Task Ordering <br> • Task Selector | • KI findings (Set of inferred tasks) <br> • User Model (demographical + current and historical data) <br> • Clinical/Functional requirements | • The most adequate content of the daily sessions |

*Table 1: Summary of the main tasks, inputs and outputs of the 4 modules that are included in the data processing layer.*

| ID | Main tasks | Configuration in restrictive scenario | Configuration in flexible scenario | Expected Planning |
|----|-----------|---------------------------------------|------------------------------------|-------------------|
| 1 | Welcome | P:100/m:-/M:-/C:1 | P:100/m:-/M:-/C:1 | Always the 1st task |
| 2 | Welcome Reminder | P:100/m:-/M:-/C:1 | P:100/m:-/M:-C:1 | Always the 1st task |
| 3 | Daily Mood Check 1 | P:100/m:-/M:-/C:3 | P:100/m:-/M:-/C:3 | Always the 2nd task |
| 4 | Daily Mood Check 2 | P:100/m:-/M:-/C:1 | P:100/m:-/M:-/C:1 | |
| 5 | Daily Mood Check 3 | P:100/m:-/M:-/C:1 | P:100/m:-/M:-/C:1 | Randomly planned between 3rd and 7th position |
| 6 | Daily Mood Check 4 | P:100/m:-/M:-/C:1 | P:100/m:-/M:-/C:1 | |
| 7 | Daily Mood Check 5 | P:100/m:-/M:-/C:1 | P:100/m:-/M:-/C:1 | |
| 8 | Select Session Type | P:80/m:-/M:-/C:2 | P:80/m:-/M:-/C:1 | Always the 8th task |
| 9 | Sleep Questionnaire | P:100/m:-/M:-/C:1 | P:20/m:-/M:-/C:1 | |
| 10 | Configure Activity Plan | P:50/m:-/M:-/C:4 | P:20/m:-/M:-/C:3 | |
| 11 | Generate Report | P:2/m:-/M:-/C:3 | P:20/m:-/M:-/C:3 | |
| 12 | Introduce Relaxation Task | P:10/m:1/M:3/C:3 | P:20/m:1/M:7/C:3 | |
| 13 | Negative Thoughts Activity Introduction | P:60/m:3/M:7/C:2 | P:20/m:1/M:7/C:2 | Position determined by the restrictions in the scenario |
| 14 | Ask For Planned Activities | P:10/m:2/M:3/C:10 | P:20/m:1/M:7/C:10 | |
| 15 | Speech Activity | P:20/m:3/M:5/C:3 | P:20/m:1/M:7/C:3 | |
| 16 | Activity Monitoring | P:5/m:-/M:-/C:3 | P:20/m:-/M:-/C:3 | |
| 17 | Introduce PHQ-9 | P:10/m:-/M:-/C:3 | P:20/m:-/M:-/C:3 | |
| 18 | Crisis Plan | P:-/m:-/M:-/C:- | P:-/m:-/M:-/C:- | Executed immediately (interruption task) |
| 19 | Farewell | P:1/m:-/M:-/C:1 | P:1/m:-/M:-/C:1 | Always the last task |

*Table 2:* Settings used during the generation and planning of the actions –dialogue acts– for both scenarios. *This table lists (1) the name of the tasks, (2) the configurations in the two defined scenarios (P=priority, m= minExecutions, M= maxExecutions, C=number of constrains), and (3) the expected planning performed by the session planner module. Tasks that do not contain the minExecutions and maxExecutions values are only planned if they are inferred by the KI (based on the detected patient condition).*

| ID Patient | Pilot | Age | Gender | Country |
|---|---|---|---|---|
| P1.1 | 1 | 30 | Female | Romania |
| P.1.2 | 1 | 26 | Female | Romania |
| P.1.3 | 1 | 23 | Female | Romania |
| P.1.4 | 1 | 27 | Female | Romania |
| P.1.5 | 1 | 45 | Female | Spain |
| P.1.6 | 1 | 39 | Female | Spain |
| P.1.7 | 1 | 38 | Female | Spain |
| P.2.8 | 2 | 49 | Female | Spain |
| P.2.9 | 2 | 60 | Female | Spain |

*Table 3: Information about the participants involved in the evaluations with pilot. In the first pilot, seven participants were included. For the second, only 2 participated.*

| ID question | Text |
| --- | --- |
| Q1 | Adaptability in the length of the sessions was adequate |
| Q2 | Variability in the content of the sessions was adequate |

*Table 4: The two questions used in the evaluation to collect feedback about the perceived variability and*

*adaptability of the sessions*

|  | Mean duration of the session (number of planned tasks) | |
| --- | --- | --- |
|  | Flexible scenario | Restrictive scenario |
| Short session | 13.4 | 14.8 |
| Medium session | 17.4 | 19.7 |
| Long session | 24.9 | 23.3 |

*Table 5: Mean values of the distributions of the simulated sessions for each type of session and scenario*

| | Flexible scenario | | | Restrictive scenario | | |
|---|---|---|---|---|---|---|
| | Simulations | Repeated | Repeated (%) | Simulations | Repeated | Repeated (%) |
| Short | 3.306 | 782 | 23,65% | 3.263 | 873 | 26,75% |
| Medium | 3.374 | 179 | 5,31% | 3.347 | 207 | 6,18% |
| Long | 3.320 | 36 | 1,08% | 3.390 | 107 | 3,16% |
| General | 10.000 | 997 | 9,97% | 10.000 | 1.187 | 11,87% |

*Table 6: Repeated sessions in the simulations using both scenarios for each type of session*

| ID question | Pilot 1 | | | | | | | Pilot 2 | |
|---|---|---|---|---|---|---|---|---|---|
| | P1.1 | P1.2 | P1.3 | P1.4 | P1.5 | P1.6 | P1.7 | P2.1 | P2.2 |
| Q1 | +1 | +1 | -1 | -1 | +1 | -1 | -1 | +1 | +1 |
| Q2 | -1 | +1 | 0 | 0 | -1 | -1 | -1 | +1 | +1 |

*Table 7: Responses of the users (P1-P9) regarding the perceived adaptability (Q1) and variability (Q2) from both*

*Pilots. The possible answers were: (-1) non-adequate, (0) neither, and (+1) adequate.*