2013 IEEE International Conference on Healthcare Informatics

On-Demand Virtual Health Counselor for Delivering Behavior-Change Health Interventions

Reza Amini
School of Computing and Information Sciences
Florida Int'l University
Miami, Florida 33199
Email: ramin001@fiu.edu

Christine Lisetti
School of Computing and Information Sciences
Florida Int'l University
Miami, Florida 33199
Email: lisetti@cis.fiu.edu

Ugan Yasavur
School of Computing and Information Sciences
Florida Int'l University
Miami, Florida 33199
Email: uyasa001@fiu.edu

Naphtali Rishe
School of Computing and Information Sciences
Florida Int'l University
Miami, Florida 33199
Email: rishe@fiu.edu

***Abstract*—In this paper, we discuss a novel approach for the computer-delivery of Brief Motivational Interventions (BMIs) for health behavior change. We describe the basic elements of our system architecture, and focus on enabling a multimodal Embodied Conversational Agent (ECA) to deliver the health behavior change interventions *empathetically* by adapting, in real-time, its verbal and non-verbal communication messages to those of its clients. The designed empathy model integrates a cognitive component and an affective components. We then discuss the evaluation experiment that we designed and conducted to evaluate the impact of empathy model on users' experience with the empathic character. Results indicate that, in comparison with the non-empathic counselor, the empathic one is better accepted (e.g., more enjoyable, empathizing, engaging, and likable) and some users might be willing to disclose more private information (e.g., drinking habits) to the counselor endowed with empathic abilities than the one without.**

I. Introduction

Unhealthy behaviors such as excessive alcohol consumption place people at risk of serious health problems. In the cases that these unhealthy behaviors are the main reasons of the health problems, identifying and changing these behaviors can prevent many of their associated diseases. Therefore, finding ways to make people more aware of their unhealthy behavior patterns and motivate them to change the unhealthy behaviors can have great impacts on peoples' health and well-being.

Research [1], [2] shows that an effective way of creating awareness and motivation is using the computer-based systems that aim at changing unhealthy behaviors. For example, the Drinker's Check Up (DCU) system [3] uses a patient-centered counseling technique called Motivational Interviewing to help people find motivation in changing their unhealthy alcohol consumption behaviors. The DCU is reported to be able to decrease alcohol consumption by an average of 50% in a 12 month follow-up.

Among different modalities of delivering the computer-based material to the computer users (e.g., text, voice, video), the Embodied Conversational Agents (ECAs) [4] are interesting user interfaces that are envisioned to be helpful in computer-based therapy and to be able to interview patients with substance abuse problems [5]. Lisetti [5] identified the key features of designing the ECAs for therapeutic purposes as: (1) human face as the interface; (2) ethnicity concordance; (3) **empathy**; (4) user-modeling; and (5) natural language abilities.

The ECAs can deliver the information and interact with the users in both verbal and non-verbal (e.g., facial expressions and body gestures) modalities. Based on the Media Equation Theory [6], if computers display social cues to their human users, the human users respond socially to them. Therefore, enabling the computer-based behavior-change systems to deliver some social cues to the clients can improve the acceptance of the clients, engage them better, and affect the outcomes positively.

However, main concerns in a delivering face-to-face interactions with clients in the health and behavior change context are: (1) building rapport and empathizing with the clients [7] which supports the clients emotionally and help them overcome their negative affects [8], (2) engaging the clients enough to the interaction so they have motivation to continue the interaction and attend the follow-up sessions [1], [2], instead of dropping out which is a significant problem with both computer-based interventions and face-to-face interventions [9], [10].

In this research, we have developed an **Empathic On-Demand VIrtual Health Counselor (Emp-ODVIHC)** which delivers the computer-based Brief Motivational Interventions (BMIs) through an ECA (see Figure 1). The BMIs include *assessment* of target behavior patterns; normative *feedback*; and

Fig. 1: Emo-ODVIHC Amy in her office.

978-0-7695-5089-3/13 $26.00 © 2013 IEEE
DOI 10.1109/ICHI.2013.13

providing a *menu* of change options to the client depending on client's readiness. The BMIs have been evaluated as effective [11] and well-accepted by the users [12]. We have selected a BMI called Motivational Interviewing (MI) [7], which is a directive, client-centered counseling style for eliciting behavior change by helping clients to explore and resolve ambivalence.

We have compared our empathic virtual health counselor with a non-empathic health counselor in terms of the clients' acceptance (e.g., perceived ease of use, enjoyment, attitude) and perceived character features (e.g., likability, anthropomorphism).

II. Related Work

In behavior-change context, there are different ECA-delivered systems from which we introduce a few in this section, and discuss their approaches to empathize with the clients. We also discuss a few of the state of the art in computational modeling empathy for the ECAs.

The MIT FitTrack [13] is an avatar-based system which uses an ECA to investigate the ability to establish and maintain a long-term working alliance with users in a behavior-change context. Their agent creates rapport using social and empathic dialogs, politeness, and nonverbal behaviors (e.g., smile). Comparing to an equivalent agent without any deliberate social-emotional, their agent was reported as more respected, liked, and trusted.

The FitTrack was used to develop the Virtual Hospital Discharge Nurse [14] to explain written hospital discharge instructions to patients with low health literacy on their hospital beds and help to review material before discharging them from the hospital. The hospital patients with low health literacy found this system easy to use, reported satisfaction, and said they preferred receiving the discharge information from the agent over their doctor or nurse.

Schulman et al. [15] designed a conversational agent as a virtual counselor for health behavior change. They use techniques drawn from MI to enhance client motivation and confidence to change. The users reported satisfaction from using this system.

Although the avatars employed in these mentioned health systems have shown some promising acceptance by their users, they have two major limitations: (1) because they are 2D, they lack dynamic expressiveness which is a key factor of communicating facial affect [16] and essential in establishing the MI requirement of *empathic* communicative style [17]; (2) they do not have a computational model of empathy so they can adapt their behaviors to the affective states of the clients, engage the clients to the intervention, motivate them to use the system in long-term, and motivate the clients to change unhealthy behaviors.

In the state of the art of computational modeling of empathy, generally, three types of approaches are taken to model the empathy: (1) mimicking the client behaviors (called affective or parallel empathy) [18]–[22], (2) including understanding and cognition to decide about the empathic reaction (called cognitive empathy) [23]–[27], or (3) empathizing with the users with a combination of the mimicry and cognitive empathy [28], [29]. In the rest of this section, we will discuss these research, their limitation, and how we can address them.

Gonsier et al. [18] and Hegel et al. [19] modeled empathy with mimicry of the user's affective state with facial expressions of a robot and showed that the subjective performance of the robot and the user acceptance is improved.

Gratch et al. [20]–[22] simulated the rapport and the positive feedback of a virtual character (called Rapport Agent) to a human by mimicking the human's postures, and reported that the users perceived the conveyed rapport and engaged to the interaction more than a neutral agent.

Prendinger, Becker-Asano, and Boukricha [23]–[25] modeled empathy by mapping the user's affective state, which is multi-modally recognized, to positive and negative facial expression using a belief-desire-intention (BDI) component. Their users reported that an agent capable of providing cognitive empathy is perceived more intelligent and is more able to provide more appropriate reactions than a neutral agent.

Pereira et al. [26] enabled a robot to assess the affective states of a chess game player and react empathetically using facial expressions and verbal comments. They report that the robot-player friendliness is improved comparing to a neutral robot.

Huang et al. [28] enhanced the Rapport Agent and developed the Virtual Rapport 2.0, in which some of the simple behavior mimicry rules are replaced by probabilistic models of the back-channel prediction, end-of-turn (turn-taking opportunity), and affective feedback (smile) based on the data driven from a video corpora. Results show that, the mutual attention, coordination, positive emotion communication, rapport, naturalness, and back-channel prediction of the Rapport Agent are improved.

In a research by Ochs et al. [29], the agent guesses the most probable emotion of the user in a BDI-like approach, and mimics the same emotion with the corresponding emotional facial expressions. They show that such an empathic character is perceived more jovial, expressive (more natural), pleasant, warm, compassionate, and cheerful than a non-emotional agent.

Moridis et al. [30] implemented a virtual tutoring agent which provides empathy with the students during a problem solving experiment. They show that, mimicry of the emotional states in a tutoring application can reinforce the student's emotion but providing positive empathic reaction can induced negative emotions to neutral in this context.

However, there are multiple limitations that can be addressed in these approaches such as: (1) non-realtime recognition of the affective state of the subjects, which can be addressed using realtime facial expression recognizers, (2) not using the facial expressions as the most important modality in human behavioral judgment [31], which can be addressed by using highly expressive characters that are capable of expressing different facial expressions, (3) using rapport and empathy in non-emotional contexts such as gaming and story telling, which can be addressed by using the rapport and empathizing ability in health counseling context, (4) unclear mapping of the user's recognized features to the character's

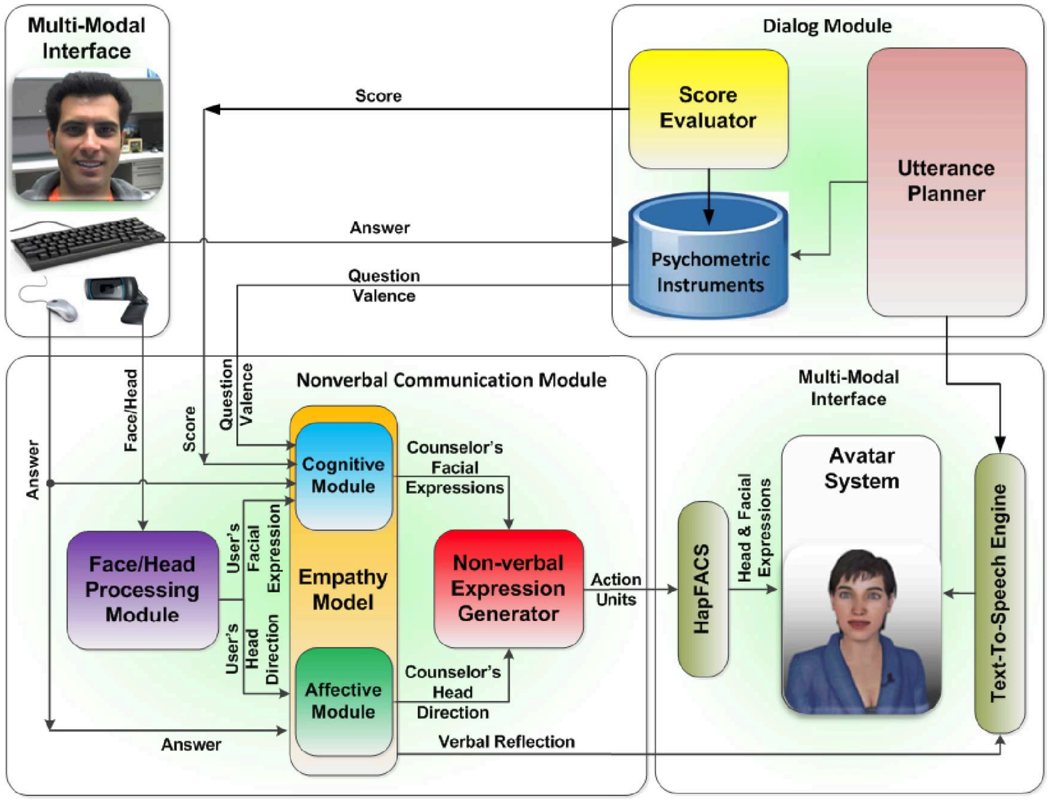


Fig. 2: ODVIC Architecture.

reactions, which can be addressed by creating an empathy model, and (5) using 2D inexpressive characters which limits conveying the nonverbal behaviors and can be addressed by using highly expressive 3D characters.

III. Health Counselor System Architecture

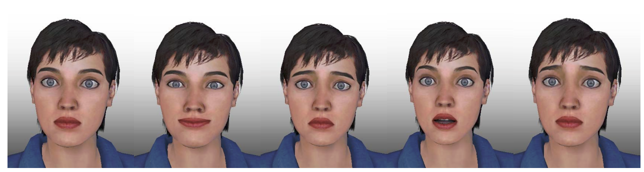
In an effort to address the limitations of current computer-based interventions discussed in Section II, we have developed an expressive 3D animated character able to empathize with the clients. Our virtual health counselor perceives the client's facial expressions and utterances during the health interaction and provides both empathic **non-verbal** expressions (emotional facial expressions, head nods, eyebrow movements), and **verbal** reflections.

A. System Overview

Our system architecture is composed of the main modules shown in Figure 2 and described in detail in [34]. The main differences between the current the previous research [34] are: (1) in the previous research empathy was modeled as mimicry of the user's emotional facial expressions, whereas, in this research, empathy is modeled both verbally and non-verbally using the two affective and cognitive modules, (2) the facial expressivity of the ECA is improved using the newly implemented module namely HapFACS (see Section III-B), and (3) we performed the character evaluation in terms of the user acceptance and perceived character features.

The **Dialog Module** evaluates and generates dialog utterances using three key components: a **Utterance Planner**, a collection of **Psychometric Instruments** (e.g. questions about drinking behavior patterns), and a **Score Evaluator**. The system.client interaction is based on a series of dialog sessions, each of which has a specific assessment goal to identify different aspects of the client's behavioral problem (e.g. dependency to alcohol, frequency of drinking).

The **Multimodal Interface** has both multimodal inputs and outputs. Multimodal inputs consist of the user's answers collected via common media (keyboard and mouse) and of the user's facial expressions captured in real-time via a webcam. The multimodal output is our 3D virtual character who speaks with lip-synch and displays a different facial expressions including the samples shown in Figure 3.



(a) (b) (c) (d) (e)

Fig. 3: Sample facial expressions: (a) neutral, (b) happy, (c) sad, (d) surprised, and (e) concerned.

The **Empathy Model** consists of an **Affective Module** and a **Cognitive Module** to communicate emphatically with the client during the behavior change intervention.

B. Modeling Empathic Communication

Discussing issues about at-risk behaviors such as heavy drinking is highly emotional for people (e.g. shame, discouragement, anger, hopefulness), and empathy and positive regard toward the client are critical therapeutic conditions to create an atmosphere of safety and acceptance where clients feel free to explore and change [35]. In MI and BMI sessions, the therapist's ability to *establish rapport and to express empathy* is crucial which can be applied by "a skillful reflective listening to clarify and amplify the [user's] own experiencing and meaning" [7].

1) Empathy Model: The *Empathy Model* emulates two kinds of empathy: *affective empathy* and *cognitive empathy*. Affective empathy involves arousal and spontaneous affective response to the perceived experience of another person [36]. Cognitive empathy involves an understanding, reasoning, and appraisal of another's experiences, combined with the capacity to communicate that understanding [37].

Consequently, we designed an empathy model for the virtual health counselor in two sub-modules of *cognitive* and *affective*. The *Empathy Model*, captures and processes user's facial expressions and head movements in real-time to assess the user's most probable affective states, then combines it with affect related information elicited from utterances to decide about the counselor's empathic responses. The *Affective Module* is responsible for fast and reactive responses such as simple verbal reflection of user's answers and head posture mimicry. The *Cognitive Module* on the other hand, is responsible for feedbacks that are more deliberative and need more thinking and decision making before expression, such as facial expressions, and head nods.

The Empathy Model uses a set of inputs to decide about the counselor's empathic behaviors:

1) *Emotional facial expressions*: facial photos are taken using the camera with the JPEG-Cam Flash/Javascript library and sent from the client computer to the face recognizer server. The face recognizer categorizes the client's emotional facial expressions into five categories of happy, sad, angry, surprised, and neutral. The face recognition engine is using the face recognition algorithm proposed in [38].
2) *Head movements*: the face recognizer returns degrees of the three possible head movements: *head yaw* (up/down), *head pitch* (left/right), and *head roll* (left/right roll).
3) *Smile*: the face recognizer returns the user's smiling status (open mouth smile) which is different from happy facial expression. The happiness is recognized from different movements of the face such as eyes, cheeks, and lips. But, smile is only the lips state.
4) *Counselor's question valence*: the counselor can expect whether her question will be pleasant or unpleasant for the client. So, the counselor uses the *role taking* mode of the empathy and puts herself in the client's shoes to guess the client's emotion in response to each question. For example, the question "Has someone been injured as a result of your drinking?" is likely to raise negative feelings while the question "Have you enjoyed the taste of wine or liquor?" may raise positive feelings.
5) *Client's answers to the counselor's questions*: the client provides an answer to each counselor's question using the mouse/keyboard. These answers show whether her/his alcohol consumption is at risk level or not.
6) *History of the client's previous answers*: after each client's answer, the **Score Evaluator** module calculates a score for the client base on her/his answers until then. Based on the user-model we proposed in our previous research [39], in different assessment sessions, this score can represent the strength of the client's dependence to alcohol, drinking risk factors, motivation to change, frequency of drinking, and consequences of drinking.

2) Affective Module: Given the above parameters, the empathy model decides which affective/cognitive empathic responses to express. The Affective Module returns a *simple* verbal feedback to each client's answer from a pool of predefined verbal feedbacks for that answer. Counselors used *verbal reflection* to create a stronger connection with the clients and create closeness and rapport. Reflection can be a repetition or rephrase of the client's response, for example, the counselor asks "How often do you have a drink containing alcohol?", the client answers "Two to three times a week", then the counselor reflects back "So, you drink at least two times a week".

Furthermore, the affective module mimics the client's head movements by mapping them to the same head movements of the counselor (head posture mimicry) to create closeness and mutual gaze with the client.

3) Cognitive Module: The *Cognitive Module* is a rule-based system which uses a set of pre-defined rules in a decision tree to decide about the next counselor's empathic reaction to the client. This module decides "what facial expression to express", "when to show head nods", and "what eyebrow expressions to show". A sample part of the used decision tree in the cognitive module is shown in Figure 4.

4) Facial Expression Generator: The *Facial Expression Generator* generates the character's facial expressions and head movements based on the Facial Action Coding System (FACS) [40], and Emotional FACS (EmFACS) [41]. FACS is a system to taxonomize human facial expressions, head, and eye movements. Movements of individual facial muscles are encoded by Action Units (AU) which are the fundamental actions of individual muscles or groups of muscles. FACS is a common standard to systematically categorize the physical expressions of the face and emotional facial expressions. EmFACS is a method for using FACS to determine the facial actions that are relevant to expressing specific emotions (happiness, sadness, fear, disgust, contempt, surprised, and anger).

The Facial Expression Generator module uses HapFACS [42], an open source[1] software that we implemented to generate facial expressions based on FACS and EmFACS. It accepts

[1] http://ascl.cis.fiu.edu/projects.html

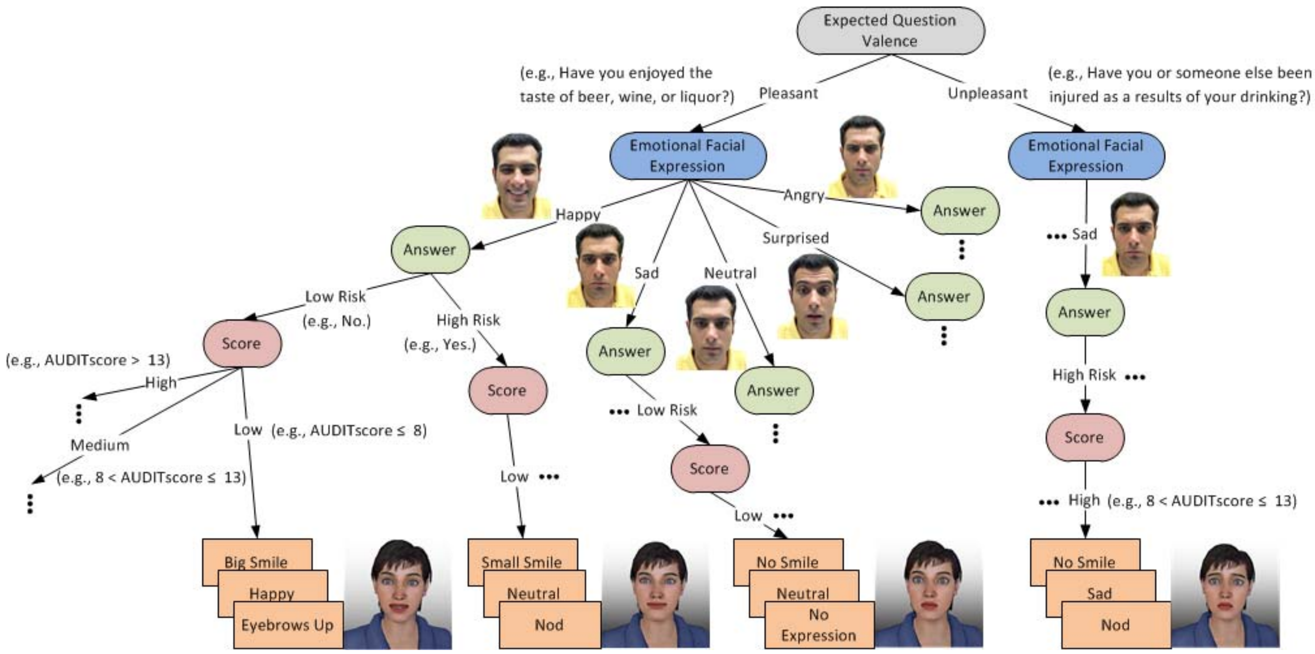


Fig. 4: A sample piece of the decision tree used in the cognitive module.

the facial expressions, head nods, and head movements that the counselor should express from the cognitive and affective modules and maps them to their corresponding AUs. Then, HapFACS generates the facial expressions, head nod, and head movements on the character face and head. For example, the sample emotional facial expressions presented in Figure 3 are generated with the following combination of AUs: (a) neutral: all AUs are deactivated, (b) happy: AU12 + AU6, (c) sad: AU1 + AU4 + AU15, (d) surprised: AU1 + AU2 + AU26, and (e) concerned: AU1 + AU4.

5) Other considerations: Lisetti [5] identified the key features of designing the ECAs for therapeutic purposes as: (1) human face as the interface; (2) natural language abilities; (3) user-modeling; (4) empathy; and (5) ethnicity concordance.

We simulate the first two features with a 3D graphical avatar well-accepted by users as documented in earlier studies [43] developed by the Haptek[2] company. We integrated the avatar with a Text-To-Speech (TTS) engine able to read text with a *natural voice* with *lip synchronization*. We use the male and female voices from the Loquendo[3] company. Furthermore, we use different mark-ups to control the avatar's TTS vocal intonation in order to increase the agent's believability and make the sound natural.

We keep a dynamic model of the user in the a User Model presented in a previous research [39], which addresses the third feature mentioned above. The virtual counselor empathizes with the client using the Empathy Model which satisfies the fourth feature mentioned above.

As shown in Figure 5, in order to address the ethnicity concordance feature we provide a set of different avatars with different *ethnicities* (e.g. black, white) in both genders to make client-counselor race and gender concordance.

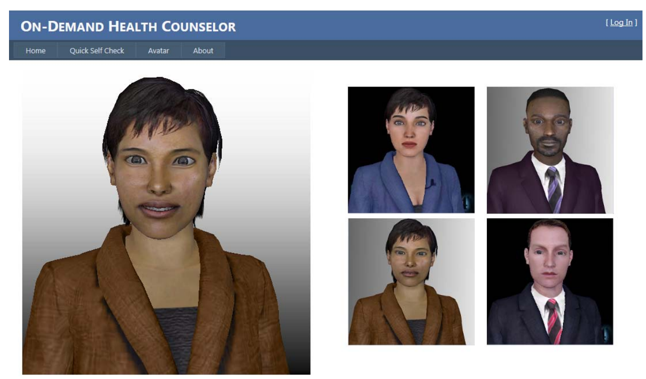


Fig. 5: Ethnicity concordance.

IV. Experiment and Evaluation

Our evaluation scheme includes the evaluation of user's acceptance of the virtual counselor and evaluation of the character's properties (e.g., likability, perceived safety).

A. Procedure

We asked the clients to attend the first session of an interview with our virtual counselor, which includes the AUDIT [44] psychometric instrument to assess the client's dependence to alcohol and frequency of drinking. AUDIT involves a set of questions about the user's drinking behaviors and a statistical feedback. For each question, the client selected an answer from a list of 3-5 answers. The clients sat in front of a computer

[2]http://www.haptek.com
[3]http://www.loquendo.com

with a camera attached to it. They had access to a keyboard and a mouse to answer to the counselor's questions. Users had the option to choose their preferred counselor's gender and ethnicity among the available characters (Caucasian, African American). The default counselor was a Caucasian female (AMY) that speaks in English.

We have implemented two conditions for the experiment:

1) **Empathic** counselor: during the counseling session, AMY reacts to the client with empathic verbal and non-verbal reactions. She expresses different emotional facial expressions (happy, sad, concerned, surprised, and neutral); head gesture (nod); big/subtle smile; head posture mimicry (pitch, yaw, roll); eyebrow movement; mutual gaze; and lip synchronized verbal reflections. Furthermore, since being polite and getting permission for pursuing the interview is an empathic technique, AMY requests for user's permission to continue during the interview.
2) **Non-empathic** counselor: AMY shows a neutral facial expression during the interview, does not empathize with the user, ignores user's changes of emotional state, and does not request for user's permission to continue the interview. Therefore, in the non-empathic condition, the Empathy Model is completely deactivated.

The clients are recruited from volunteer university students through fliers and emails. We selected the subjects who had at least one drink in the last month. From the total number of 51 subjects, 26 were assigned to the empathic counselor and 25 to the non-empathic counselor. Subjects included 37% females and 63% males. The ethnicity distribution of the participants was as 55% White, 27% Hispanic, 16% African American, and 2% Asian.

After the subjects interacted with the system, we debriefed them using an after-experiment questionnaire about the acceptance and performance of the counselor.

B. *Questionnaire*

We compiled a modification of the two questionnaires developed by Heernik et al. [45] and Bartneck et al. [46]. Then, we used the resulting questionnaire in an online user self-report survey to debrief the users after using our system.

Heerink's model evaluates the **users' acceptance** of assisting social artificial agents. Heerink designed 13 constructs each of which is represented by multiple statements. Users replied to these statements on a 5-point Likert scale (-2 for strongly disagree to +2 for strongly agree). We used 10 of those constructs each of which includes 1 to 5 statements. Here, we provide the definition of each construct and one example of the used statements:

- Attitude (ATT): positive/negative feelings about the appliance of the technology (e.g. I think it's a good idea to use the counselor.)
- Intention to Use (ITU): outspoken intention to use the system over a longer period in time (e.g. I think I'll use the system again.)
- Perceived Enjoyment (PENJ): feeling of pleasure associated by user with use of the system (e.g. I enjoyed participating in this session with the counselor.)
- Perceived Ease of Use (PEOU): degree to which the user believes using the system would be free of effort (e.g. I found the counselor easy to use.)
- Perceived Sociability (PS): perceived ability of the system to perform sociable behavior (e.g. I think the counselor is empathizing with me.)
- Perceived Usefulness (PU): degree to which a person believes using the system would enhance his or her daily activities (e.g. I think the counselor can help me.)
- Social Presence (SP): experience of sensing a social entity when interacting with the system (e.g. When interacting with the counselor, I felt like I'm talking to a real person.)
- trust (TRUST): belief that the system performs with personal integrity and reliability (e.g. I would trust the counselor if it gave me advice.)
- Anxiety (ANX): evoking anxious or emotional reactions when using the system (e.g. I was afraid to make mistakes during the interview.)
- Social Influence (SI): user's perception of how people who are important to him think about him using the system (e.g. I am comfortable to disclose information about my drinking to the counselor.)

Bartneck et al. [46] have defined another questionnaire model called "Godspeed" including key concepts of human-computer interaction with the following definitions:

- Anthropomorphism (ANT): attribution of a human form, characteristics, or behavior to non-human concepts such as robots, computers, and animals (e.g. I rate the counselor as Machine-like/Human-like.)
- Likability (LIKE): degree to which the agent evokes empathic or sympathetic feelings of the user (e.g. I rate my impression as Dislike/Like.)
- Animacy (ANIM): degree to which a computer agent is lifelike and can involve users emotionally (e.g. I rate the counselor as Dead/Alive.)
- Perceived Intelligence (PI): user's perception of how the agent is intelligent (e.g. I rate the counselor as Unintelligent/Intelligent.)
- Perceived Safety (PSA): user's perception of the level of danger, and her/his level of comfort during the use (e.g. During the interaction I was Agitated/Calm.)

Users answered to the statements in this questionnaire in a 5-point Likert type scale (-2 to +2).

V. Results and Discussion

A. *Results about the User's Experience*

The total of 56 statements categorized in 15 classes were asked from the users. The clients' answers were analyzed using the Mantel-Haenszel Chi-Square statistical method (with degree of freedom $df = 1$) which involves (1) assigning scores to the response levels, (2) forming means, and (3) examining location shifts of the means across the levels of the responses.

We followed a null hypothesis: counselors with different levels of empathizing abilities (empathic vs. non-empathic) have the same effects on the users. A common significance threshold (i.e., alpha) value in the chi-square analysis is 5%. Therefore, under the assumption of the null-hypothesis, p values less than 0.05 reject the null-hypothesis.

Also, we compared the mean values of the same statements in the two experimental conditions to calculate the possible improvement/deterioration of them upon each other.

1) Attitude: Results show significant differences in terms of attitude between the empathic and non-empathic conditions ($X^2 = 5.56, p < 0.05$). This result indicates that adding the empathizing ability to the character affects the attitude to use the system and the clients expect a human-like system to be empathic. This result confirms a previous research by Nguyen and Masthoff [47].

The positive mean values of empathic ($mean = 0.78, stdev = 0.9$) and non-empathic ($mean = 0.31, stdev = 1.05$) versions indicate that the clients have positive attitude toward the system regardless of the empathizing ability of the counselor. However, the mean value comparison shows that the clients have 11.81% more attitude to use the empathic counselor than the non-empathic counselor.

2) Intention to Use: Results show significant differences in terms of intention to use between the empathic and non-empathic conditions ($X^2 = 6.41, p < 0.05$). Therefore, adding the empathizing ability to the character affects the intention to use the system. This result supports the previous result that "the clients expect a human-like system to be empathic".

The positive mean values of empathic ($mean = 0.80, stdev = 0.89$) and non-empathic ($mean = 0.12, stdev = 0.89$) versions show that the clients have positive intention to use an avatar-based counselor. This result confirms the results of our previous research [43] in which 74% of the clients reported a positive intention to use the avatar-based system. The mean value comparison shows that the clients have 17.12% more intention to use the empathic counselor than the non-empathic counselor.

3) Perceived Enjoyment: Results show significant differences in terms of perceived enjoyment between the empathic and non-empathic conditions ($X^2 = 24.40, p < 0.05$). This result shows that adding the empathizing ability to the character affects the user's perceived enjoyment.

Again, it shows that the clients expect a human-like system to be empathic. The positive mean values of empathic ($mean = 0.99, stdev = 0.63$) and non-empathic ($mean = 0.31, stdev = 0.97$) versions indicate that the clients perceived the system positively enjoyable regardless of the empathizing ability of the counselor. However, the mean value comparison shows that the clients enjoyed the empathic version 17.11% more than the non-empathic one.

4) Perceived Ease of Use: Results show significant differences between the empathic and non-empathic conditions ($X^2 = 0.52, p < 0.05$) in terms of the perceived ease of use. This result indicates that adding the empathizing ability to the character affects the perceived ease of use of the system.

The positive mean values of empathic ($mean = 0.84, stdev = 1.24$) and non-empathic ($mean = 0.96, stdev = 1.27$) versions indicate that the clients perceived both of the version easy to use. However, comparison of these mean values show that enabling the character to empathize with the users complicates the use of the system.

5) Perceived Sociability: Results show significant differences between the empathic and non-empathic conditions ($X^2 = 36.57, p < 0.05$) in terms of perceived sociability. This result shows that adding the empathizing ability to the character affects the perceived social abilities of the character.

Statements in the Perceived Sociability category debrief the clients about the empathizing, understanding, and social abilities of the counselor. Therefore, the positive mean value of empathic counselor ($mean = 0.80, stdev = 0.87$) indicates that the clients perceived it empathizing, understanding, nice and sociable. On the other hand, negative mean value of the non-empathic version ($mean = -0.07, stdev = 0.97$) indicate that the clients perceived the non-empathic 21.68% less sociable than the empathic version.

6) Perceived Usefulness: Results show significant differences between the empathic and non-empathic conditions ($X^2 = 10.13, p < 0.05$). Therefore, the clients perceived the same usefulness level from non-empathic version. This result indicates that adding the empathizing ability to the character affects the perceived usefulness of the system.

The positive mean values of empathic ($mean = 0.68, stdev = 0.88$) and non-empathic ($mean = 0.02, stdev = 1.08$) versions indicate that the clients perceived the system positively useful regardless of the empathizing ability of the counselor. However, the mean value comparison shows that the clients think that the empathic counselor is 16.52% more useful than the non-empathic one.

7) Social Presence: Results show significant differences between the empathic and non-empathic conditions ($X^2 = 25.15, p < 0.05$) in terms of social presence. This result shows that adding the empathizing ability to the character affects the perceived social presence of the character.

The positive mean value of empathic ($mean = 0.21, stdev = 1.07$) indicates that the clients sense a social entity when interacting with the empathic counselor. But, negative mean value of non-empathic ($mean = -0.57, stdev = 0.99$) shows that the clients do not have this sense when interacting with the non-empathic version. The empathic counselor makes 19.43% improvement over the non-empathic version.

8) Trust: Results show significant differences between the empathic and non-empathic conditions ($X^2 = 13.01, p < 0.05$) in terms of trust. This result indicates that adding the empathizing ability to the character affects the perceived trust level of the character by the user.

The positive mean value of empathic version ($mean = 0.51, stdev = 1.07$) indicates that the clients can trust the empathic counselor. But, negative mean value of non-empathic counselor ($mean = -0.07, stdev = 1.01$) shows that it is not trustful enough for the clients. The mean value comparison shows that the empathic counselor is 14.31% more trustful than the non-empathic one.

9) Anxiety: Results show no significant differences between the empathic and non-empathic conditions ($X^2 =$

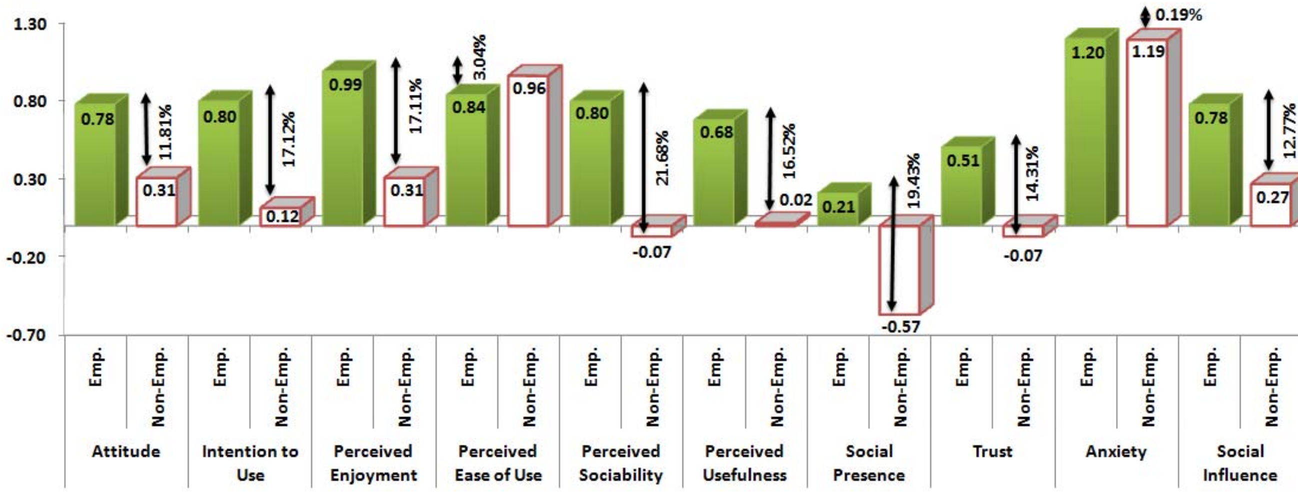


Fig. 6: Mean value comparison of experimental conditions for user acceptance features.

$0.003, p > 0.05$). This indicates that the clients feel the same anxiety level while using the two versions of the system.

The positive mean values of empathic ($mean = 1.2, stdev = 0.87$) and non-empathic ($mean = 1.19, stdev = 1.02$) versions indicate that non of the two counselor versions evoke anxious while interacting with the clients. Also, the empathizing ability of the counselor does not provide a major improvement over the non-empathic version in terms of anxiety reduction.

10) Social Influence: Results show significant differences between the empathic and non-empathic conditions ($X^2 = 5.53, p < 0.05$) in terms of social influence. This result shows that adding the empathizing ability to the character affects the possible social influences of the character.

The positive mean values of the empathic ($mean = 0.78, stdev = 1.03$) and the non-empathic ($mean = 0.27, stdev = 1.10$) versions show that the two versions have positive social influence on the clients independent of the empathizing ability. However, the empathic version is perceived 12.77% more socially influential on the users.

Figure 6 shows the mean value comparison of the two experimental conditions for the user acceptance features described above.

11)Anthropomorphism: Results show that there are significant differences between the empathic and non-empathic counselors ($X^2 = 27.42, p < 0.05$) in terms of anthropomorphism. This result indicates that adding the empathizing ability to the character affects the anthropomorphism level of the character.

The positive mean value of the empathic version ($mean = 0.28, stdev = 1.05$) indicates that the counselor was positively perceived anthropomorphic by the clients. On the other hand, the negative mean value of the non-empathic version ($mean = -0.47, stdev = 1.10$) indicates that the non-empathic version is perceived not so anthropomorphic and is is perceived 18.73% less anthropomorphic than the empathic version.

12)Likability: Results show significant differences between the empathic and non-empathic conditions ($X^2 = 21.51, p < 0.05$) in terms of likability. This result shows that adding the empathizing ability to the character affects the likability of the character by the user.

The positive mean value of empathic ($mean = 1.29, stdev = 0.64$) and non-empathic ($mean = 0.85, stdev = 0.78$) versions indicates that the clients liked both of the versions. However, the empathic version is 10.85% more likable than the non-empathic.

13) Animacy: Results show that there are significant differences between the empathic and non-empathic counselors ($X^2 = 28.59, p < 0.05$) in terms of their animacy, which indicates that adding the empathizing ability to the character affects the animacy level of the character.

The positive mean value of the empathic version ($mean = 0.68, stdev = 0.98$) indicates that the counselor was perceived well animated. On the other hand, the negative mean value of the non-empathic version ($mean = -0.11, stdev = 1.21$) indicates that the non-empathic version is not perceived so well animated and it is perceived 19.69% less animated than the empathic version.

14) Perceived Intelligence: Results show significant differences between the empathic and non-empathic conditions ($X^2 = 18.76, p < 0.05$) in terms of perceived intelligence, which shows that adding the empathizing ability to the character affects the perceived intelligence of the character by the user.

The positive mean values of the empathic ($mean = 0.93, stdev = 0.74$) and the non-empathic ($mean = 0.42, stdev = 1.04$) versions indicates that the clients perceived both of the versions intelligent. However, comparison shows that, the empathic version is perceived 12.82% more intelligent than the non-empathic version.

15) Perceived Safety: Results show significant differences between the empathic and non-empathic conditions ($X^2 =$

$11.44, p < 0.05$) in term of the level of perceived comfort/danger during the system use. This result indicates that adding the empathizing ability to the character affects the perceived safety level of the character.

The positive mean values of empathic ($mean = 1.39, stdev = 0.95$) and non-empathic ($mean = 0.79, stdev = 1.11$) versions indicate that the clients feel comfortable when using both of the versions of the system. However, the empathic version is 14.79% safer than the non-empathic one.

Figure 7 shows the mean value comparison of the two experimental conditions for the character features described above.

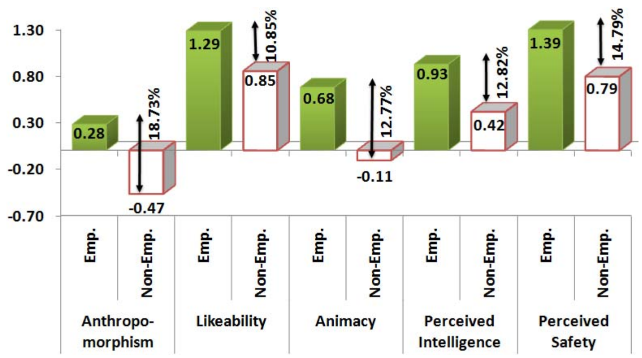


Fig. 7: Mean value comparison of experimental conditions for the character features.

VI. Conclusions and Future Directions

In this article we described the design, implementation, and evaluation of an empathic virtual health counselor who can deliver an evidence-based Brief Motivational Intervention (BMI) for behavior change on alcohol consumption. Our system is modular and can easily be adapted to other target behaviors such as overeating and lack of exercise. Users' overall acceptance of the system over a number of dimensions regarding the impact of the empathic communication of the character indicates that this novel modality of delivery for behavior change intervention has a significant impact in terms of users' motivation to continue to use such systems. Moreover, our results showed that people expect a system represented by a human-like character to be empathic.

Although the head movement mimicry of the affective module was completely implemented and integrated in the current research, it was turned off during the experiment. So, in our future studies we will turn the head gesture mimicry and evaluate its effects on the user's perception. Furthermore, we would like to compare the empathic, and non-empathic versions with a totally textual version in which no character is used to deliver the interaction. So, we will be able to reason about the effects of the presence of a character in the interaction (either empathic or non-empathic). Also, as discussed the current Empathy Model decision making is rule-based, which needs psychological expertise to define the rules. We aim to generate computational models of the non-verbal gestures for the counselor. This can be performed using machine learning from the annotated video corpora of human counselor-client interactions.

Acknowledgment

This material is based in part upon work supported by the National Science Foundation under Grant Nos. CNS-0821345, CNS-1126619, HRD-0833093, IIP-0829576, CNS-1057661, IIS-1052625, CNS-0959985, OISE-1157372, IIP-1237818, IIP-1215201, IIP-1230661, IIP-1026265, IIP-1058606, IIS-1213026, OISE-0730065, CCF-0938045, CNS-0747038, CNS-1018262, CCF-0937964.

References

[1] M. Vernon, "A review of computer-based alcohol problem services designed for the general public," *Journal of substance abuse treatment*, vol. 38, no. 3, pp. 203–211, 2010.

[2] D. B. Portnoy, L. a. J. Scott-Sheldon, B. T. Johnson, and M. P. Carey, "Computer-delivered interventions for health promotion and behavioral risk reduction: a meta-analysis of 75 randomized controlled trials, 1988-2007," *Preventive medicine*, vol. 47, no. 1, pp. 3–16, Jul. 2008.

[3] R. K. Hester, D. D. Squires, and H. D. Delaney, "The Drinker's Check-up: 12-month outcomes of a controlled clinical trial of a stand-alone software program for problem drinkers," *Journal of Substance Abuse Treatment*, vol. 28, no. 2, pp. 159–169, 2005.

[4] J. Cassell, J. Sullivan, S. Prevost, and E. F. Churchill, "Embodied Conversational Agents," *Social Psychology*, vol. 40, no. 1, pp. 26–36, 2000.

[5] C. L. Lisetti, "Embodied Conversational Agents for Psychotherapy," in *Proceedings of the CHI 2008 Conference Workshop on Technology in Mental Health*. ACM, 2008, pp. 1–12.

[6] B. Reeves and C. Nass, *The Media Equation: How People Treat Computers, Television, and New Media Like Real People and Places*. New York, NY: University of Chicago Press, 1996.

[7] W. R. Miller and S. Rollnick, Eds., *Motivational interviewing: 2nd Ed. Preparing people to change addictive behavior.* New York: Guilford Press., 2002.

[8] J. O. Greene and B. R. Burleson, *Handbook of Communication and Social Interaction Skills*, J. O. Greene and B. R. Burleson, Eds. Lawrence Erlbaum Associates, Inc., Publishers, 2003.

[9] T. L. Dunn, L. M. Casey, J. Sheffield, P. Newcombe, and A. B. Chang, "Dropout from computer-based interventions for children and adolescents with chronic health conditions." *Journal of health psychology*, vol. 17, no. 3, pp. 429–42, Apr. 2012.

[10] M. Wierzbicki and G. Pekarik, "A meta-analysis of psychotherapy dropout." *Professional Psychology Research and Practice*, vol. 24, no. 2, pp. 190–195, 1993.

[11] B. M. Bewick, K. Trusler, M. Barkham, A. J. Hill, J. Cahill, and B. Mulhern, "The effectiveness of web-based interventions designed to decrease alcohol consumption–a systematic review." *Preventive medicine*, vol. 47, no. 1, pp. 17–26, Jul. 2008.

[12] A. Cunningham, J.A., Humphreys, K., & Koski-Jannes, "Providing personalized assessment feedback for problem drinking on the Internet." *33rd annual convention of the Association for the Advancement of Behavior Therapy, Toronto.*, 1999.

[13] T. W. Bickmore and Rosalind W. Picard, "Establishing and maintaining long-term human-computer relationships," *ACM Transactions on Computer-Human Interaction (TOCHI)*, vol. 12, no. 2, pp. 617–638, 2005.

[14] T. W. Bickmore, L. M. Pfeifer, and B. W. Jack, "Taking the Time to Care : Empowering Low Health Literacy Hospital Patients with Virtual Nurse Agents," in *Proceedings of the 27th international ACM conference on Human factors in computing systems (CHI'09)*. New York: ACM, 2009, pp. 1265–1274.

[15] D. Schulman, T. W. Bickmore, and C. L. Sidner, "An Intelligent Conversational Agent for Promoting Long-Term Health Behavior Change using Motivational Interviewing," in *Association for the Advancement of Artificial Intelligence (AAAI) Spring Symposium Series*. Association for the Advancement of Artificial Intelligence (www.aaai.org), 2011, pp. 61–64.

[16] S. M. Ehrlich, D. J. Schiano, and K. Sheridan, "Communicating Facial Affect : It's Not the Realism, It's the Motion," in *Proceedings of ACM CHI 2000 Conference on Human Factors in Computing Systems*. NY: ACM, 2000, pp. 252–253.

[17] W. R. Miller and S. Rollnick, *Motivational Interviewing: Preparing People for Change*, 2nd ed. New York: Guilford Press, 2002, vol. 2nd, no. 1.

[18] B. Gonsior, S. Sosnowski, C. Mayer, J. Blume, B. Radig, D. Wollherr, and K. Kuhnlenz, "Improving aspects of empathy and subjective performance for HRI through mirroring facial expressions," in *RO-MAN, 20th IEEE International Symposium on Robot and Human Interactive Communication*. Atlanta, GA, USA: IEEE, 2011, pp. 350–356.

[19] F. Hegel, T. Spexard, B. Wrede, G. Horstmann, and T. Vogt, "Playing a different imitation game: Interaction with an Empathic Android Robot," in *Humanoid Robots, 2006 6th IEEE-RAS International Conference on*. IEEE, 2006, pp. 56–61.

[20] J. Gratch and S. C. Marsella, "A domain-independent framework for modeling emotion," *Cognitive Systems Research*, vol. 5, no. 4, pp. 269–306, 2004.

[21] J. Gratch, N. Wang, and A. Okhmatovskaia, "Can virtual humans be more engaging than real ones?" in *Proceedings of the 12th international conference on Human-computer interaction: intelligent multimodal interaction environments, HCI'07*. Chamonix, France: Springer-Verlag Berlin, Heidelber, 2007.

[22] N. Wang and J. Gratch, "Rapport and facial expression," in *3rd International Conference on Affective Computing and Intelligent Interaction and Workshops, 2009. ACII 2009*. Amsterdam: IEEE, Sep. 2009, pp. 1–6.

[23] H. Prendinger and C. Becker-Asano, "A Study in User's Physiological Response to an Empathic Interface Agent," *International Journal of Humanoid*, vol. 3, no. 3, pp. 371–391, 2006.

[24] H. Boukricha and C. Becker-Asano, "Simulating empathy for the virtual human max," in *Proceedings of the 2nd Workshop at KI2007 on Emotion and Computing Current Research and Future Impact*, Dirk Reichardt and P. Levi, Eds., Osnabrück, Germany, 2007, pp. 23–28.

[25] C. Becker-Asano and I. Wachsmuth, "Affective computing with primary and secondary emotions in a virtual human," *Autonomous Agents and Multi-Agent Systems*, vol. 20, no. 1, pp. 32–49, May 2009.

[26] A. Pereira, I. Leite, S. Mascarenhas, C. Martinho, and A. Paiva, "Using empathy to improve human-robot relationships," *Human-Robot Personal Relationships*, vol. LNICST 59, pp. 130–138, 2011.

[27] S. Rodrigues and S. Mascarenhas, " I can feel it too !: Emergent empathic reactions between synthetic characters," in *Affective Computing and Intelligent Interaction and Workshops (ACII'2009)*. Porto, Portugal: IEEE, 2009, pp. 1–7.

[28] L. Huang, L.-p. Morency, and J. Gratch, "Virtual Rapport 2.0," in *Proceedings of the 11th international conference on Intelligent virtual agents (IVA'11)*. Reykjaavik, Iceland: Springer-Verlag Berlin, Heidelberg, 2011, pp. 68–79.

[29] M. Ochs, D. Sadek, and C. Pelachaud, "A formal model of emotions for an empathic rational dialog agent," *Autonomous Agents and Multi-Agent Systems*, Nov. 2010.

[30] C. N. Moridis, A. A. Economides, and S. Member, "Affective Learning : Empathetic Agents with Emotional Facial and Tone of Voice Expressions," *IEEE Transactions on Affective Computing*, vol. 3, no. 3, pp. 260–272, 2012.

[31] N. Ambady and R. Rosenthal, "Thin slices of expressive behavior as predictors of interpersonal consequences: A meta-analysis." *Psychological Bulletin*, vol. 111, no. 2, pp. 256–274, 1992.

[32] D. Schulman, T. Bickmore, and C. Sidner, "An Intelligent Conversational Agent for Promoting Long-term Health Behavior Change using Motivational Interviewing," in *2011 AAAI Spring Symp. Series*, 2011, pp. 61–64.

[33] C. R. Rogers, "A theory of therapy, personality and interpersonal relationships as developed in the client-centered framework," in *Psychology: the Study of a Science*, S. Koch, Ed. New York: McGraw-Hill, 1959, vol. 3, ch. 3, pp. 184–256.

[34] C. L. Lisetti, "10 Advantages of using Avatars in Patient-Centered Computer-based Interventions for Behavior Change," in *IHI2012 International Health Informatics Sysmposium*, Miami, FLorida, 2012.

[35] W. R. Miller and G. S. Rose, "Toward a Theory of Motivational Interviewing," *American Psychologist*, vol. 64, no. 6, pp. 527–537, 2009.

[36] L. Wispé, "History of the concept of empathy," in *Empathy and its development*, N. Einsenberg and J. Strayer, Eds. Cambridge University Press, 1987, ch. 2, pp. 17–37.

[37] M. Hojat, *Empathy in patient care: antecedents, development, measurement, and outcomes*. New York, NY: Springer, 2007.

[38] L. Wolf, T. Hassner, and Y. Taigman, "Effective Unconstrained Face Recognition by Combining Multiple Descriptors and Learned Background Statistics." *IEEE transactions on pattern analysis and machine intelligence*, vol. 33, pp. 1–13, Dec. 2010.

[39] U. Yasavur, R. Amini, and C. L. Lisetti, "User Modeling for Pervasive Alcohol Intervention Systems," in *First International Workshop on Recommendation Technologies for Lifestyle Change 2012 (LIFESTYLE 2012)*, Dublin, Ireland, 2012, pp. 29–34.

[40] P. Ekman and W. V. Freisen, *Facial Action Coding System: A Technique for the Measurement of Facial Movement*, C. P. Press, Ed. Consulting Psychologists Press, 1978.

[41] W. V. Friesen and P. Ekman, "EMFACS-7: Emotional Facial Action Coding System," *Unpublished manuscript, University of California at San Francisco*, 1983.

[42] R. Amini, U. Yasavur, and C. L. Lisetti, "HapFACS 1.0: Software/API for Generating FACS-Based Facial Expressions," in *Proceedings of the ACM 3rd International Symposium on Facial Analysis and Animation (FAA'12)*. Vienna, AUSTRIA: ACM Press, 2012.

[43] C. L. Lisetti, U. Yasavur, C. D. Leon, R. Amini, and N. Rishe, "Building an On-demand Avatar-based Health Intervention for Behavior Change," in *Proceeding of FLAIRS'2012 Association for the Advancement of Artificial Intelligence (www.aaai.org)*, no. Mi, Miami, FL, US, 2012.

[44] T. F. Babor, J. C. Higgins-Biddle, J. B. Saunders, and M. G. Monteiro, *AUDIT: The Alcohol Use Disorders Identification Test. Guidelines for use in primary health care*, 2nd ed. World Health Organization, Department of Mental Health and Substance Dependence, 2001.

[45] M. Heerink, B. Krose, V. Evers, and B. Wielinga, "Measuring acceptance of an assistive social robot: a suggested toolkit," in *Robot and Human Interactive Communication, 2009. RO-MAN 2009. The 18th IEEE International Symposium on*. IEEE, 2009, pp. 528–533.

[46] C. Bartneck, D. Kulic, and E. Croft, "Measuring the anthropomorphism, animacy, likeability, perceived intelligence and perceived safety of robots," in *Proceedings of the Metrics for Human-Robot Interaction Workshop in affiliation with the 3rd ACM/IEEE International Conference on Human-Robot Interaction (HRI 2008), Technical Report 471*, vol. 471. Amsterdam: University of Hertfordshire, 2008, pp. 37–44.

[47] H. Nguyen and J. Masthoff, "Designing empathic computers: the effect of multimodal empathic feedback using animated agent," in *Proceedings of the 4th International Conference on Persuasive Technology (Persuasive09)*. Claremont, California, USA: ACM, 2009.