Virtual Coach for Mindfulness Meditation Training

Eva Hudlicka
Psychometrix Associates, Inc.
1805 Azalea Drive, Blacksburg, VA , US
hudlicka@ieee.org

Abstract

The past decade has witnessed an increasing interest in the use of virtual coaches in healthcare. This paper describes a virtual coach to provide mindfulness meditation training, and the coaching support necessary to begin a regular practice. The coach is implemented as an embodied conversational character, and provides mindfulness training and coaching support via a web-based application. The coach is represented as a female character, capable of showing a variety of affective and conversational expressions, and interacts with the user via a mixed-initiative, text-based, natural language dialogue. The coach adapts both its facial expressions and the dialogue content to the user’s learning needs and motivational state. Findings from a pilot evaluation study indicate that the coach-based training is more effective in helping users establish a regular practice than self-administered training via written and audio materials. The paper concludes with an analysis of the coach features that contribute to these results, discussion of key challenges in affect-adaptive coaching, and plans for future work.

Introduction and Objectives

The past decade has witnessed an increased interest in the use of embodied conversational characters (ECAs) (Cassell et al., 2000). In a number of applications these characters function as virtual coaches, to provide assistance, training, or coaching support across a variety of contexts, or as virtual humans in a variety of training contexts. Examples include health behavior coaching; e.g., virtual coach Laura’s exercise coaching (Bickmore & Giorgino, 2006), virtual patients used in training contexts (e.g., virtual Alzheimer’s patient (Green, 2009), advisors for patients with specific medical conditions (e.g., agent Greta who provides advice about eating disorders (de Rosis et al., 2003), and helpers and trainers for caregivers (Johnson et al., 2004).

In this paper we describe a virtual coach that provides mindfulness meditation training, and the coaching support necessary to begin a regular practice. The coach is implemented as an embodied conversational character, and provides mindfulness training and coaching support via a web-based application. The coach is represented as a female character, capable of showing a variety of affective and conversational expressions, and interacts with the user via a mixed-initiative, text-based, natural language dialogue. The coach adapts both its facial expressions and the dialogue content to the user’s learning needs and motivational state. Findings from a pilot evaluation study indicate that the coach-based training is more effective in helping users establish a regular practice than self-administered training via written and audio materials. The paper is organized as follows. First, we provide background information about mindfulness, embodied conversational agents and affect-adaptive interaction. Next, we describe the virtual mindfulness coach prototype. Next, we discuss an empirical evaluation study and the findings regarding the coach’s effectiveness. The paper concludes with a discussion of technical challenges in creating effective virtual coaches, focusing on their multi-modal interaction capabilities and social and affective realism.

Background and Related Work

Below we provide a brief summary of the relevant background information regarding mindfulness meditation, and embodied conversational agents and virtual coaches.

Mindfulness Meditation Mindfulness is best described as a state of awareness characterized by “full attention to, and awareness of, the internal and external experience of the present moment” (Chambers et al., 2008, p.304), and a non-judgmental attitude towards whatever thoughts, images, feelings or sensations enter awareness. A typical mindfulness practice lasts about 20 minutes. The meditator sits in a comfortable position, and begins the session with a concentration practice, by focusing his/her attention on some fixed, concrete entity, usually the breath, followed by mindfulness practice proper, where
the meditator observes whatever emerges into consciousness, with full, non-judgmental awareness. Regular practice usually involves 20-30 minutes, done once or twice daily.

Regular practice of mindfulness meditation is associated with number of health benefits. Evidence from existing studies indicates that mindfulness practice enhances health-related quality of life, affecting both physical and psychological symptoms (Lazar, 2005; Reibel et al., 2001). Benefits of mindfulness meditation have been demonstrated in stress reduction (Shapiro et al., 1998), pain reduction (Kabat-Zinn, 1985), enhanced immune responses (Davidson et al., 2003), reduction of symptoms in anxiety disorders (Evans et al., 2008), prevention of relapse in major depression (Teasdale et al., 2000), improvement in a subjective sense of well-being (Brown & Ryan, 2003), and improvements in cognitive functions (Chambers et al., 2008).

In spite of these benefits, the relative ease of learning mindfulness techniques, and no costs associated with their practice, use of mindfulness meditation remains limited. This is due to several factors, including lack of adequately trained healthcare providers, lack of access to teachers and training programs, cost of training programs, misconceptions about the methods involved, and difficulties with establishing a regular practice.

The virtual mindfulness coach prototype described here was developed to address this problem, and to provide increased access to mindfulness training and coaching. The coach mindfulness training content is based on a widely-used mindfulness training program, the Mindfulness Based Stress Reduction (MBSR) (Kabat-Zinn, 1990). The mindfulness meditation training consists of 4 lessons, with one lesson administered each week during the 4-week training period. Each lesson focuses on a specific topic, and the 4 lessons together are designed to provide a solid introduction to mindfulness techniques, address common problems beginners encounter (e.g., restlessness, boredom), and suggest how to integrate mindfulness into daily life.

**Embodied Conversational Agents and Virtual Coaches**

Virtual coaches use the emerging technologies of embodied conversational agents (ECA’s) (Cassell et al., 2000), and relational pedagogical agents (Bickmore, 2003). ECA’s are animated virtual characters, displayed on a computer or a mobile device screen. ECA’s play the roles of teachers, mentors, advisors, social companions, and, increasingly, of virtual coaches (Prendinger & Ishizuka, 2004; Hayes-Roth, 2004; Rickel et al., 2002; Clarebout et al., 2002; Johnson et al., 2000). The use of ECA’s requires minimal or no training, due to their capability to engage in natural interaction with humans through dialogue and non-verbal expression. The latter includes facial expressions, gaze, and gestures, which together help control conversation flow and augment the ECA’s visual and behavioral realism. Relational pedagogical agents represent a subset of ECA’s, designed both to train a particular subject or skill, and to develop the type of a longer-term relationship with the user that is necessary to facilitate coaching (Bickmore et al., 2005).

The use of ECA’s and relational pedagogical agents has recently begun to be explored in healthcare settings, where these virtual entities act as stand-ins for humans, and as trainers, coaches and counselors. ECA’s acting as human stand-ins are being explored in the training of psychiatrists in diagnostic interviewing skills (Kenny et al., 2008), as virtual coaches to help individuals establish and maintain exercise programs (Bickmore, 2003), and to adopt healthy eating habits (deRosis et al., 2003), and as virtual counselors to provide problem-solving skills training and emotional support for caregivers (Johnson et al., 2004).

Both the visual appearance and the multi-modal interaction capabilities of these characters vary greatly. The embodiments range from cartoonish, animated characters or faces, to fully-articulated full bodies. The interaction capabilities range from text-based, multiple-choice user input and text-based character output, to speech recognition and synthetic speech. The virtual characters also vary in the degree to which they can adapt to the user’s knowledge, affective and motivational state, and display affective and social realism.

**Virtual Mindfulness Training Coach Prototype**

The Virtual Mindfulness Coach is implemented as an embodied conversational agent, “Chris”, who guides the student through the training material, and provides supportive coaching to help students establish a regular mindfulness practice (refer to figure 1).

Chris engages multi-modal verbal and non-verbal interaction with the user. Chris is represented by a drawn image of a face, which can show a variety of affective and conversational expressions (e.g., happy, sad, concerned, confused). The verbal interaction is conducted via a mixed-initiative, natural-language, text-based dialogue (refer to figure 2). The coach can also support limited non-verbal interaction, by displaying affective and conversational expressions that match the content of the dialogue, provide appropriate affective reaction to the user’s utterances, and display conversational expressions that aim to communicate the coach’s mental state (e.g., understanding, confusion).

These capabilities enable the coach to provide customized coaching and affect-adaptive feedback to the user; e.g., smile when the user reports success, express concern and verbal encouragement when the user reports a problem.

The coach is able to support several types of pedagogical strategies necessary for effective training of mindfulness meditation, and the coaching required to initiate and maintain a regular practice: didactic (to convey the conceptual basis of mindfulness, information
User-input is highlighted in yellow. Coach follows a lesson plan (red text), but allows for interruptions from the student to follow specific topics of interest.
about the techniques, and their health benefits); experiential learning of mindfulness techniques (via the guided meditation recordings, narrated by an experienced mindfulness teacher); and supportive and coaching interaction, designed to maintain student motivation via empathic dialogue and customized, concrete advice.

Coach Chris acts as a guide through the mindful lesson materials, which are displayed primarily via segments of text, and augmented by Chris’ introductions and elaborations, that emphasize material Chris thinks is of particular interest to the user, based on previous dialogue. Chris delivers the mindfulness training in 4 lessons, which follow the same structure; introduction and overview, guided meditation, main lesson, Q&A, adaptive segment, homework, good-bye. Chris also offers coaching, during which she prompts student to discuss specific problems, and offers suggestions and advice, including pointers to existing didactic material, which she displays for the student.

The coach thus guides the student through the didactic lesson material, including recorded guided meditations, and provides customized advice about meditation practice, based on the student’s expressed concerns. For example, if the student repeatedly asks about meditation and its effects on stress, the coach follows-up with questions about the student’s stress levels, and provides customized suggestions for practice.

Multiple Types of Interactions Chris interacts with the user via text-based natural language (see figure 2 for an example of a dialogue segment). The dialogue between Chris and the user is displayed in a dedicated pane on the coach user interface (see the “Dialogue with Coach “Chris” pane in figure 1). Chris can engage in several types of dialogue with the user: didactic, relational and adaptive/motivational. Chris can also answer the user’s free-form questions.

During the training segment, Chris engages primarily in didactic interaction, as she guides the user through the training materials, by introducing the different lesson segments, and displaying the didactic material in a dedicated pane of the coach UI (see “Main Text Window” in figure 1). Throughout the lesson, Chris checks-in with the user, to ensure understanding and motivation. For example, after displaying the material for the weekly lesson, Chris asks the user whether s/he understood everything, and offers to provide additional information for any items that require elaboration. Periodically, Chris engages in relational interaction to promote relationship building with the user, and user engagement. For example, at the beginning of each lesson, Chris asks the user how their meditation went the previous week, and expresses joy or concern, depending on the user’s answer.

To facilitate the adaptive/motivational interaction, Chris collects information about the user throughout the interaction, and periodically asks specific questions (e.g., “Did your meditation go ok last week?” “Were you able to meditate 5 times?”). The information provided by the student is used to gradually construct a user model, which is then used to support adaptive/motivational dialogue, by customizing responses and questions to the student’s knowledge and motivational state. For example, Chris tracks how many questions the student asks about a particular topic, how well the practice is going, what specific problems the student is encountering, etc.

Based on this information, Chris then asks a series of ‘adaptive questions’ at the end of each lesson, to provide customized interaction. For example, if the student asked a number of conceptual questions about mindfulness (e.g., “What’s the theory of mindfulness?”; “How does mindfulness work?”), Chris offers to provide additional information. If the user asks repeated questions about stress (e.g., “Is mindfulness good for stress?” “Will meditating help me feel less stressed?”), Chris first attempts to determine whether the student is feeling stressed, and then offers additional material on mindfulness and stress. Chris frames the questions in a tentative manner, to allow for the possibility that her understanding of the student’s motivational state is incorrect; e.g., “You asked several questions about mindfulness and stress. Is this something of particular concern to you?”. If the student answers ‘Yes’, Chris responds: “Many people practice mindfulness for stress reduction. Would you like to hear additional information bout mindfulness and stress reduction?”

Natural Language Understanding and Dialogue Management The user has two options for natural language dialogue with coach Chris: free-form natural language input, and multiple-choice based input. In the free form input mode, the user can interrupt the coach at any point and make a comment, or ask a question related to mindfulness practice. The natural language understanding component is currently relatively simple, and is implemented via keyword and template matching. User input is scanned and matched against a list of stored keyword templates, until a match is found. Chris then responds with the phrase associated with that template. If the user input is not understood, Chris expresses disappointment (“I am sorry, but I didn’t understand what you meant”) and displays an expression of confusion or concern. Chris also suggests that the user can type the question again, so that it can be added to her database.

This approach has several limitations, including lack of robustness (any new output must be explicitly included in the template database), the need for extensive template ‘tuning’, limited domain of discourse, and a relatively high rate of user utterances that are not understood or misunderstood. Nevertheless, by appropriately structuring of the coach dialogue, and providing information to the users about the coach’s limited NL understanding capabilities, effective dialogue between the coach and the student can take place. Figure 2 illustrates a segment of the coach-student dialogue, highlighting the mixed-initiative capabilities.
The second mode of user input, via multiple-choice based options, was implemented to address the relatively high frequency of user utterances that were not understood by the coach. This enhancement was based on the success of multiple-choice based input in health coach Laura (Bickmore, 2003). This user input alternative is emphasized during the coaching sequence, where the coach asks the student to select a particular problem encountered, or a goal on which to focus, from a set of multiple choice options (refer to “Multiple Choice user input in figure 1). For example, the following are some of the options provided when the coach asks the user whether s/he has any concerns about the practice: Can’t find the time; Meditation is too boring; Keep falling asleep; Feel too restless; Pain makes it difficult to sit still; Not sure meditation is the right thing for me; etc. Based on the user’s response, further, context-sensitive options are then displayed. The user’s responses are echoed in the dialogue pane, so that a continuous dialogue is displayed for the user (refer to figure 1). With careful and thorough analysis of the domain and the student’s likely needs, adequate sets of multiple-choice options can be developed to minimize the likelihood that a particular option will not be available.

This approach to augmenting the free-form text input proved to be effective, and reduced the users’ frustration with Chris’ inability to understand all of the student utterances. In fact, after this enhancement was implemented, the users reported that Chris’ ability to understand free-form natural language input improved, even though this component was not changed.

**Affective Modeling and Adaptation**

An important component of training and coaching interaction is the coach’s ability to adapt to the student’s knowledge and motivational state. An effective coach should be able to customize its feedback to the student’s specific knowledge needs (provide additional information or correct misconceptions), affective state (express empathy and understanding when the student is sad), and motivational state (provide support and encouragement when the student is lacking in self-confidence, and challenges when the student is becoming bored). A core component of this capability is the coach’s ability to recognize and adapt to the student’s emotional state.

To implement this type of adaptation, coach Chris collects information about the student from questions asked, comments, and answers to Chris’ questions. Data collected include frequency of questions on a particular topic (e.g., stress, health benefits, restlessness), frequency of positive vs. negative comments, student’s assessment of the overall experience, number of questions asked, etc.

Together, this information comprises the user model and the affective user model. Chris uses this information to deepen the relationship with the user (via customized relational dialogue), to provide adaptive training by providing specific information to enhance the user’s knowledge of the training content, and to enhance engagement and motivation by providing appropriate emotional support and encouragement.

**Implementation**

The first version of the coach was implemented using proprietary software developed by LifeLike Solutions, Inc. A version 2 coach was later developed in-house, to provide us with the capability to modify and augment the software to introduce new functionalities (e.g., the multiple-choice based user input). The coach is currently deployed via the web. The user interface is defined in HTML, with embedded Javascript supporting the parsing of user input necessary to support NL understanding and student-coach interaction, and coach affective expressions. External data files define the keyword template-phrase pairs, the facial expressions, and the didactic material. The affective user model is currently stored in local variables via cookies.

**Evaluation Study**

The coach’s effectiveness was evaluated in a 7-week study, designed to assess the coach’s ability to provide mindfulness training (weeks 1-4), and to support students in establishing a regular meditation practice (weeks 5-7). During the 4 week training period, one lesson was administered per week. During the 3 week coaching period, the participants were free to interact with the coach (for the experimental group), or read the manual (for the control group), as needed.

A between-subject design was used, with two conditions: an experimental coach group (N=16), using the Virtual Mindfulness Coach, and a control group (N=16), using written and audio materials (CD with guided meditations). Participants in both groups were asked to meditate 5x/week, for 20 minutes each time.

The participants (N=32; 75% female) were recruited from a local medical school (25%), and the general community (75%). Mean age was 38. Participants were pseudo-randomly assigned to the two conditions, to provide gender- and age-balanced samples. Control group mean age was 40; control group mean age was 36.

Data were collected via web-administered surveys, using SurveyMonkey.com. The format of most questions was a 5-point Likert scale (Strongly Agree, Agree, Neutral, Disagree, Strongly Disagree). All surveys included opportunities for free-form feedback.

Weekly surveys were administered to track the frequency of meditation, and the participants’ overall experience with the training and their meditation practice, including any benefits experienced, and any difficulties encountered in establishing a regular practice. At 5 weeks, two tests were administered to assess the participants’ knowledge of mindfulness concepts and techniques. During the final week, a survey was administered assessing the participants’ sense of self-efficacy regarding their ability to continue their mindfulness practice, as well as their stage of change within the transtheoretical model of change (Prochaska &
DiClemente, 2005). Self-efficacy was assessed via a single (Likert scale) survey item: “I feel confident that I can continue my practice after this study is finished”. To assess the stage of change, the participants were asked to indicate their assessment of 3 negative and 2 positive aspects of their overall experience with meditation. The negative aspects were assessed via the following items: “Meditation takes a lot of time”, “Meditation requires a lot of effort”, and “It is difficult to find the time”. The positive aspects were assessed via the following items: “There are tangible benefits associated with meditation”, “The benefits I have derived from meditation are worth the time and effort”.

Participants in the experimental group filled out additional surveys assessing specific features and functionalities of the Virtual Mindfulness Coach, and their experience with the coach (at weeks 1, 3, 6 and 7). These also included baseline and final surveys regarding the coach user interface. The focus was on assessing the coach’s affective and social realism, ability to effectively answer questions, and its adaptive and relational capabilities.

The results of the study indicate that the coach provides more effective training and coaching than a self-administered program using written and audio materials. Specifically, the coach group participants practiced more frequently (4.5 days/week vs. 3.2), and for longer time (19’ vs. 16’), than the controls. This effect was particularly evident during the 3-week follow-up period (frequency: 4.3 days/week coach group vs. 2.7 control; length: 18.5’ coach group vs. 13.6’ control). The coach group participants found the experience more rewarding, enjoyable, beneficial, and engaging than the controls. The coach group also demonstrated a general increased sense of involvement, as demonstrated by the lower number of dropouts than the control group (3 vs. 6), and higher frequency of answering the surveys.

In addition, the coach group participants appeared to be in more advanced stages of change, in terms of the transtheoretical model of behavior change, and felt more confident in their ability to maintain a regular meditation practice, compared with the controls. All of these findings were statistically significant (p < .05). The conceptual and practical knowledge about mindfulness meditation was equal for both groups.

These results provide support for the hypothesis that virtual coach-based training and coaching is not only effective for mindfulness meditation training, but is more effective in helping students establish a regular practice, than the use of written and audio materials. The fact that the results were statistically significant in spite of the small number of participants is particularly encouraging. However, a more extensive study is necessary to confirm these findings.

Interestingly, the assessment of several coach features and functionalities was not as positive as we had anticipated, or as the meditation frequency data above would suggest. The users felt neutral to mildly positive about the coach’s ability to provide customized feedback and address questions or difficulties (.3 on a Likert scale ranging from -2 to +2). They were slightly more positive about the coach’s ability to provide support and encouragement (.9). The assessments of the coach's ability to adapting non-verbal expressions to the student’s needs, and a feeling of ‘personal connection’ with the coach, were negative (-.6 and -.5, respectively).

Discussion

The study data suggest that the most successful coach feature was its ability to provide customized feedback and support, in an interactive manner. This was the case in spite of the negative assessments of the coach’s ability to conduct free-form, natural language dialogue, and the negative assessment of the coach’s affective and social realism.

Several specific hypotheses should be further explored to explain these findings. First, it is likely that the negative assessment of the coach’s social and affective realism was due to two factors: the frequent misunderstood user input, and the appearance of the coach’s face. Additional experiments are necessary to identify the role of these two factors in the overall sense of the coach’s affective and social realism. It is possible that the participants experienced a phenomenon referred to as the ‘uncanny valley’ (MacDorman, 2005; Mori 1970), where the sense of overall believability of synthetic agent or robot negatively correlates with the degree of visual realism of the agent / robot appearance. In other words, more ‘cartoonish’ agents are perceived as more believable and affectively realistic than more visually-realistic agents. The reason for this apparently counterintuitive finding is that the unconscious criteria used to evaluate believability and affective realism shift as the appearance of the agent becomes more human-like. In effect, for more cartoonish characters, our expectations are lower. Once the synthetic agent begins to resemble human appearance, our evaluation criteria become more stringent, we begin to expect human-like realism and effectiveness in verbal and non-verbal interaction, and when these expectations are not met, we feel that the character is not believable. The role of the ‘uncanny valley’ phenomenon in the coach-participant interaction can be further explored, by varying the degree of visual realism of the coach’s face, to determine an optimal level of realism to establish an engaging relationship.

Since preference for faces are highly idiosyncratic, it is also possible that the reason for the negative assessments of the coach’s believability and non-verbal interaction was that the participants simply did not like the face of coach Chris, and would have felt a greater sense of believability, social realism and engagement with a different face. Future versions of the coach will explore this possibility, by providing the users with a set of alternative faces, including photographs, and including arbitrary images provided by the user. For example, a user
would be able to upload an image of his/her favorite teacher.

Finally, there is also the possibility that the coach face did not play a major role in the effectiveness of the interaction, and that the dialogue itself, the linguistic coaching support, and the ability to rapidly select relevant material (e.g., ability to browse relevant segments of content), were responsible for the coach’s effectiveness. This hypothesis will be evaluated in future studies, by exploring the relative effectiveness of an embodied conversational agent coach (such as the one used for the prototype) vs. a coach capable of dialogue but not represented by a face. Such an alternative was already briefly explored in the version 2 coach, where the participants had the option to replace the image of coach Chris by an image of an nature scene. A number of participants stated a preference for the nature scene over the coach face image.

Several additional hypotheses regarding the specific coach features that correlate with an overall positive (or negative) perception will be investigated in the future. For example, the addition of the multiple-choice based input in version 2 of the coach not only enhanced the overall user satisfaction with the coach, and increased the perception that the coach was able to understand user input, but appeared to result in a perception that the free form dialogue was more robust in the version 2 coach, when in fact there were no changes in the free-form dialogue.

Identification of the specific features of an embodied conversational agent that contribute to its believability, and to the user’s level of engagement, is an active area of research. Future studies will focus on a more extensive evaluation of individual coach features with respect to their effect on the overall coach believability, effectiveness, and a sense of engagement experienced by the user.

**Summary and Conclusions**

We described a virtual coach prototype for mindfulness meditation training and coaching. The coach is implemented as an embodied conversational character “Chris”. To support multi-modal, adaptive interaction with the user, Chris can display a variety of facial expressions, and modify her dialogue, to appropriately respond to user’s input and knowledge and affective state. Chris can engage in didactic, relational and supportive coaching interactions with the student, via a mixed-initiative, text-based natural language dialogue.

Results of an evaluation study of Chris’ effectiveness with 32 participants indicates that the virtual coach is superior to a self-administered mindfulness course in providing mindfulness meditation training and coaching, and results in more frequent meditation. A more extensive study with a larger sample size is necessary to confirm these preliminary findings.

These results are encouraging, and provide evidence that even a relatively simple coach structure (in terms of NL understanding, dialogue management, affective realism, affective user modeling and affect-adaptation) can have significant impact on the degree of student engagement during training, and on learning outcomes.

Future work will focus on enhancing the coach’s NL understanding and generation capabilities and dialogue management, and the affective and social realism of the coach’s embodiment and interaction, via more complex user emotion recognition and affective modeling (Hudlicka, 1999; Hudlicka, 2002; Hudlicka & McNeese, 2002). In addition, several specific hypotheses will be explored in future studies to identify the key factors that contribute to student engagement and learning.

**Acknowledgements**

This work for performed under an NIH SBIR grant 1R43 AT005397-01, awarded to Psychometrix Associates by the National Center for Complementary and Alternative Medicine. The author would like to acknowledge the support of Dr. John Glowa of NIH, and thank the following individuals who contributed to this research: Kevin Griffin, who developed the mindfulness training materials; Jonathan Klein, who designed and developed the coach’s face image and expressions, and developed the coach user interface; George Roberts, who designed and implemented the coach Javascript parsers; and Drs. Noreen Klein and Marion Reynolds of Virginia Tech, who provided statistical consultations.

**References**


