Int. J. Human–Computer Studies 159 (2022) 102745

Contents lists available at ScienceDirect

International Journal of Human - Computer Studies

journal homepage: www.elsevier.com/locate/ijhcs

Investigating the use of speech-based conversational agents for life coaching

Laura Aymerich-Franch, PhD *, Iliana Ferrer, PhD

Pompeu Fabra University, Roc Boronat, 138, 08018 Barcelona

ARTICLE INFO

Keywords:
Conversational agents
Coaching
Psychological well-being
Positive technology
Goal achievement

ABSTRACT

Life coaching can contribute to goal attainment, quality of life, and psychological well-being enhancement. We explored the capacity of a speech-based conversational agent coach (CAC) to deliver a coaching program for goal achievement in two studies. Participants showed a significant increase in personal growth initiative (PGI) after completing the program both in the pilot and the main study. Participants in the main study additionally reported a significant increase in life satisfaction (SLS) and a significant decrease in negative affect (PANAS-N). Usability of the application, satisfaction with the coaching program, and adoption intention were rated positively in both studies. The results suggest that working on goal achievement with the CAC had a positive impact on the psychological well-being of the participants. The study provides an empirically-validated approach for automated coaching interventions and highlights the potential of conversational agents for delivering life coaching.

1. Introduction

Life coaching is a result-oriented process in which a coach facilitates the enhancement of life experience and goal attainment in personal and professional life (Grant, 2003). Coaching focuses on boosting personal development and promoting the positive aspects of the human being, rather than on psychopathology (Grant, 2003). Previous studies have found that coaching contributes to enhancing psychological well-being, quality of life, goal achievement, and resilience (Grant, 2003; Grant et al., 2009; Green et al., 2006).

Due to their complexity, coaching sessions are typically conducted by professional human coaches. Here, we wanted to explore whether a speech-based conversational agent could be implemented for this purpose.

Conversational agents are systems that communicate with users using natural language, either written or oral (Diederich et al., 2019). Speech-based conversational agents use natural language processing to interpret human speech and synthesized voices to communicate with humans. A widely expanded form of speech-based conversational agents are voice assistants (e.g., Siri or Google Assistant), which run on purpose-built speakers, smartphones, and other devices. It is estimated that nearly half of the adult population in the United States uses them (Pew Research Center, 2017).

Voice assistants are currently used in everyday life for functions such as music playback, content reading, or obtaining the weather forecast (Hoy, 2018). They add some other features usually developed by third-party developers that expand their basic functionalities by interfacing via voice with services, apps, and devices, all interconnected through cloud services (Hoy, 2018).

The actual range of functions available to the final user on these devices is still very limited. Another disadvantage is the lack of a common metric to evaluate the replies against human judgement for conversational agents (Merdivan et al., 2020). However, thanks to their rapid expansion, affordability, ability to understand and respond using natural language, and being perceived as engaging and natural (Pew Research Center, 2017), voice assistants, and speech-based conversational agents in general, present a great potential for a wide range of yet unexplored uses.

One of the areas in which speech-based conversational agents could expand is for health-care and well-being related applications (ter Stal et al., 2020). E-health applications have been highlighted as promising tools that can encourage the adoption of healthy behaviors (Kreps and Neuhauser, 2010) and palliate access challenges to health services (Whitten et al., 2001). In this regard, previous works have already examined potential applications of speech-based conversational agents to promote physical (Albaina et al., 2009; Watson et al., 2012) and mental (Hudlicka, 2013; Ly et al., 2017) health.

Some studies refer to conversational agents as coaches when they are

* Corresponding author.
E-mail address: laura.aymerich@gmail.com (L. Aymerich-Franch).

https://doi.org/10.1016/j.ijhcs.2021.102745
Received 11 March 2021; Received in revised form 10 November 2021; Accepted 20 November 2021
Available online 27 November 2021
1071-5819/© 2021 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

developed for well-being related purposes (Albaina et al., 2009; Hudlicka, 2013; Watson et al., 2012). However, they generally do so in the sense of trainer, tutor, advisor, or recommender, rather than in the sense of a conversational agent that adopts the role of a life coach and delivers a full coaching program based on techniques from the field of Coaching. In fact, the lack of a precise understanding of what constitutes an e-coaching system and how it differs from other types of approaches has been previously highlighted (Kamphorst, 2017).

Coaching has available a series of techniques for self-development and well-being improvement which need to be recognized as inherent to the discipline (Kamphorst, 2017; Whitmore, 2010). The role of the coach is not to teach pupils, give suggestions, or advise clients, as it might be the case with other approaches such as tutoring, mentoring, or consulting. The primary aim of coaching is to accompany the coachee in forming well-designed goals and developing an effective action plan, to stimulate ideas and action, and to ensure that the goals are consistent with the coachee's main life values and interests (Ives, 2008).

Here, we introduce a speech-based conversational agent life coach (hereafter, CAC) that delivers a complete coaching program for goal achievement based on well-established coaching techniques. The program is aimed at the adult, non-clinical population. The CAC is inspired by the foundations of Positive Technology (Riva et al., 2012), which promotes the development of emerging technologies to improve the quality of life and well-being of people.

2. The coaching approach

A key strategy of the coaching practice is the asking approach (Stoltzfus, 2008; Whitmore, 2010). Coaching questions are intended for boosting reflection and leading to meaningful change. They are characterized by being open questions, and generally start with *what, when, who, where,* or *how/how much*. The questions relating to *why* are typically avoided in coaching practice for being considered to evoke defensive reactions rather than information responses or facts (Whitmore, 2010).

The process of asking questions is usually structured around a conversational model that allows advancing towards setting an agenda, defining what to work on, creating meaningful goals, developing options, addressing obstacles, and taking action (Stoltzfus, 2008). The coaching conversation can be also supported with other exercises and techniques.

Many coaches conduct some initial assessments to gain an overall view of where the client stands in life at that given moment to help them identify what areas in their life need more work. The contributions of the Quality of Life Therapy (QOLT) approach (Frisch, 2006, 2016) are particularly useful in this regard. According to QOLT, the value or importance assigned to a particular area of life is key to understanding life satisfaction (Frisch, 2006). Satisfaction in highly valued areas, the ones that the person cares about the most, has a great influence on overall life satisfaction (Frisch, 2006). For that, it is important to put energy into the areas that are most relevant to the person. Highly valued areas are not universal, as they are relative to each person. The *happiness pie technique* (Frisch, 2006) is an exercise aimed at identifying and reordering personalized priorities in this respect. By comparing two pies, one that represents the current time and energy dedicated to each of the main areas of the person's life, and another that represents the ideal time and energy dedicated to these areas, the individual is able to reflect on potential imbalances between the two versions and realize what areas need to be prioritized (Frisch, 2006).

In addition, the importance of setting goals in alignment with the person's values has been highlighted by multiple coaching approaches (Cox et al., 2014). Values define what is most important to the person and can act as a driving force in individuals' goals and actions. However, values are deeply ingrained assumptions, for which people are not generally consciously aware of them (Stoltzfus, 2008). Values discovery as a preceding exercise to goal setting is an important process to become in touch with who the person is and what they truly value in life. Including values discovery in the coaching process provides a more holistic approach to personal development which should ultimately contribute to enhanced well-being (Grant and Cavanagh, 2010). Values discovery is best approached by examining and selecting specific values for each of the main areas of life (Stoltzfus, 2008).

Concerning goal-setting, the SMART criteria (Doran, 1981) are one of the most widely used tools in coaching for this purpose (Stoltzfus, 2008; Whitmore, 2010). While several versions exist regarding the meaning of the mnemonic acronym, a commonly utilized version is that SMART stands for *Specific, Measurable, Achievable, Realistic,* and *Time--based*. In particular, *specific* refers to the fact that the goal states specifically what the person wants to accomplish. *Measurable* means that the goal contains some accountable aspects and has indicators, so that it is possible to evaluate if the goal is being attained. *Achievable* is related to the fact that it is in the power of the person to achieve the goal, and will not depend on anyone else. *Realistic* means that the person can realistically achieve their goal, and *time-based* is related to the fact that the person needs to think of a time frame, regarding when the results will be achieved.

Finally, a common practice to support goal attainment in coaching are visualization techniques. Research suggests that imagining one's future success can contribute to improved performance (Ruvolo and Markus, 1992). The aim of visualization exercises is to promote goal-directed behavior by increasing the person's expectation for success (Alberts and Poole, 2019). A typical form of visualization involves listening to a script of a successful experience in which the individual is invited to close their eyes, relax, and imagine the situation described from a first-person perspective (Aymerich-Franch and Bailenson, 2014; Ayres and Hopf, 1993).

2.1. Coaching intervention program for the CAC application

The coaching program used in our studies was developed *ad hoc* for the CAC application. Guided by the CAC, participants decided the area of life in which they wanted to work, found their core values for that specific area, built a goal based on the SMART criteria for goal setting, and put the goal into practice. The experimenters did not intervene in the coaching process in any way and the participants had complete freedom to decide about the goal they wanted to work on and implement.

The coaching program used in the main study had a total length of three sessions, which were delivered weekly for three weeks. The version initially developed and used in the pilot study contained four sessions, consisting of an introductory session and three coaching sessions. For the final version of the coaching program, used in the main study, the introductory session and the first coaching session were merged into a single session (see *6.3.Results* of the Pilot Study).

Sessions were individual and unfolded in the following manner. Session 1 started with a self-introduction of the CAC, which presented itself as "NORIKA, your virtual coach". Subsequently, there was an introduction to the coaching program and to the functioning of the app, the initial well-being questionnaires (see *5.2. Measures*) were completed, and then, there was an introduction to the basic dynamics of coaching. Following this, the participants were guided through a variant of the happiness pie exercise (Frisch, 2006). For that, participants were requested to draw two pies, one that represented the current time and energy dedicated to each of the main areas of their life and another that represented the ideal time and energy dedicated to these areas. Some examples of areas were given, but the participants were free to list the areas that applied to their own life, with no pre-established categories. Then, the CAC asked a series of questions (e.g., what area/s of the cake really need/s changes to see an enhancement in your well-being?) that invited to reflection and ultimately allowed the participants to decide in what area of life they wanted to work during the program. After that, the CAC helped the participants identify their core values for that specific area. In the last part of the session, the CAC guided the participants

through a series of questions to identify the changes that were needed in that particular area and what could be done to achieve those changes. This was done in alignment with the core values of the participant. The participants then chose one of the particular ideas that had listed and re-wrote it following specific indications in line with the general coaching practice for goal setting (e.g., write the idea in positive, not in negative). The resulting goal was kept as the potential goal in which the participants were going to work on in the coming sessions. At the end of the session, the participants stated the extent to which they were committed to achieving their goal.

In session 2, the SMART criteria for goal setting were introduced and the participants were guided towards transforming the goal they chose in the previous session into a goal that met the SMART criteria. Participants were also given the option of choosing a new action from the list if they were not satisfied with the goal they had from the previous session. If they did so, they were guided again through the same process conducted in session 1 until they had the goal ready, and then applied the SMART criteria. As in the previous session, the participants stated the extent to which they were committed to achieving their goal at the end of the session.

In session 3, participants initially answered to what extent they had achieved their goal so far. Then, the goal was revised and reframed if needed following the five dimensions of the SMART criteria. Then, a visualization process was conducted in which the participants were invited to imagine they had already achieved their goal. After thoroughly experiencing this state, they were asked to realize all the steps they did to reach that point. Once the visualization concluded, they were guided towards writing as specifically as possible what they were going to do to achieve their goal, as well as how, when, and with whose support, in accordance to what they saw in the visualization. To conclude the session, the participants were asked again about their level of commitment to their goal, and the final well-being questionnaires were completed (see *5.2. Measures*).

Table 1 summarizes the content of the sessions.

3. Technical characteristics of the CAC

A natively developed app for iOS was built for implementing the CAC. The iTunes file sharing was used to upload the files containing the different coaching sessions, the accompanying visual material, and to deliver and collect data for the questionnaires (see *5.2. Measures*). The application was adapted for iPhone and iPad. For the current studies, a 5.5-inch screen iPhone was used.

During the session, the app used the native iOS text-to-speech to guide the participants through the session by voice. To facilitate user interaction, the application used the native speech-to-text abilities so it could discern among a series of pre-set possible answers which were used to decide how the session developed among a set of possible options.

For the coaching program, the range of possible answers on each decision point was limited to either one or two. One possible answer was used in some reflection and written activities for the participants to indicate that they had completed the activity and were ready to proceed. This was indicated with the word "next". This word was also used in more complex answers as a final word so that the app understood that the participants had finished their intervention. Two possible answers were used in the dialog to respond to yes-no type questions. An example of the conversation between the CAC and the participant is as follows:

CAC: Now that your goal is set, we will do a visualization exercise to help you achieve your goal. Are you ready?
PARTICIPANT: Yes / No
CAC (if Yes): Excellent. Let's continue
CAC (if No): No problem, you can pause and resume when you are.

The participants were able to interact with the CAC verbally and through a touchscreen. The coaching sessions were principally voice-based, with some accompanying visual material (e.g., illustration images).

The interface was designed to be very simple and easy to use. Also, the non-embodied coach design was aimed at focusing exclusively on the guiding voice of the CAC and to avoid unnecessary distractions during the session. A simple turquoise big static circle with a smaller white inner circle was displayed in the middle of the screen during the sessions, for the entire session, except when some visual material was presented.

Additionally, there were two buttons on the bottom side of the screen. One to pause the session, on the right side, and another to go back and repeat the last explanation by the CAC, on the left side. The *Menu* button was reserved for the experimenters to select the coaching session. Fig. 1 shows the interface with which participants interacted during the sessions.

For the questionnaires, each item was presented on the screen and the participant selected the answer on a Likert scale using the touchscreen. The answers were encrypted and stored locally on the device. The files containing the results were later downloaded on a computer and decrypted for the analysis.

The coaching program was initially delivered in English in the pilot study. For the main study, the coaching program was translated into Spanish to facilitate the recruitment of local volunteers.

For the English version, a female-like and a male-like American English-speaking voices were made available to the participants to choose their virtual coach among the text-to-speech voices available in iOS. When the program started, the voice, regardless of which one was selected, introduced itself as NORIKA. For the Spanish version, a female-like voice was used for being the only one available in Castilian Spanish.

A voice sample of the CAC (English version, female-like) is available in the supplementary material (see SI – CAC).

Table 1
The three-week coaching program with the CAC.

Session	Aim	Contents
1	Defining a goal[1]	- Introduction to coaching program / app functioning - Self-evaluation satisfaction areas of life - Determining a specific area of life for improvement - Identifying core values for that area - Defining a goal in line with the core values
2	Working towards the goal I	- Introduction to the SMART coaching model - Working towards achieving the goal based on the SMART model
3	Working towards the goal II	- Working towards achieving the goal based on the SMART model (II) - Visualization exercise

[1] In the pilot study, Session 1 was divided into two sessions: introduction to the program (Session 1) and defining a goal (Session 2).

Fig. 1. Interface of the coaching application that was displayed during the sessions.

4. Hypotheses

We expected that completing the coaching program with the CAC would positively contribute to enhancing the psychological well-being of the participants. In particular, the following hypotheses were defined:

H_1. The participants will experience an increase in personal growth after completing the coaching program with the CAC, compared to before.
H_2. The participants will experience an increase in life satisfaction after completing the coaching program with the CAC, compared to before.
H_3. The participants will experience an increase in positive affect (H_{3a}) and a decrease in negative affect (H_{3b}) after completing the coaching program with the CAC, compared to before.

We additionally evaluated the usability and adoption intention of the CAC, and satisfaction with the coaching program (see Holzinger et al. (2008) for a revision of usability attributes). We defined the following research questions regarding these variables:

RQ_1. How do participants evaluate the usability of the CAC?
RQ_2. What is the participants' willingness to adopt the CAC in real life?
RQ_3. What is the participants' level of satisfaction with the coaching program?

5. General method

5.1. Overview

We conducted a pilot study and the main study to examine the effectiveness of the CAC. Both studies utilized a within-subject design and the sessions were individual. The participants read the information of the study prior to their participation and signed a consent form to participate. The study received ethical approval from the University where the study was conducted. The pilot study took place within the University premises. Due to the COVID-19 pandemic outbreak, the procedure was modified to add new safety protocols and to guarantee that the main study could be conducted in a regular manner (i.e., weekly sessions) in the current situation (e.g., intermittent closures of the University buildings, limitations of free movement, etc.). For this reason, and after approval by the ethical committee, the main study was conducted in the participants' homes.

5.2. Measures

To assess the effectiveness of the coaching application to enhance psychological well-being, life satisfaction, and personal growth, the following measures were used to collect data before and after the study:

Personal Growth Initiative Scale - PGI (Robitschek, 1998). This scale is used in coaching to evaluate the active, intentional engagement in changing and developing as a person. The scale contains both cognitive and behavioral components and an overall orientation toward change (Magyar-Moe, 2009). The scale contained 9 items (e.g., *I know how to change specific things that I want to change in my life; I have a good sense of where I am headed in my life; I have a specific action plan to help me reach my goals*) rated on a 7-point scale (1- strongly disagree – 7 - strongly agree). The reliability[1] of the scale was $\alpha = 0.89$ (before) and $\alpha = 0.94$ (after).

Satisfaction with Life Scale – SLS (Diener et al., 1985). This scale is used to measure global life satisfaction. We adapted the original scale and retained 4 items (e.g., *I am completely satisfied with my life*) which were rated on a 7-point scale (1- strongly disagree – 7 - strongly agree). It has been suggested that this scale can be utilized to better understand general levels of life satisfaction as the person cognitively (versus emotionally) reflect on their experiences in the world (Magyar-Moe, 2009). Reliability of the scale was $\alpha = 0.76$ (before) and $\alpha = 0.88$ (after).

Positive and Negative Affect Scale - PANAS (Watson et al., 1988). This scale consists of two 10-item scales to measure positive and negative affect. Participants rated the items on a 7-point scale (1 - not at all – 7 - very strongly). It has been suggested that combining the use of the PANAS and the Satisfaction with Life Scale (Diener et al., 1985) can offer a good overview of where the person stands in terms of an overall level of emotional well-being (Magyar-Moe, 2009). Reliability of the scale was $\alpha = 0.86$ (positive, before), $\alpha = 0.95$ (positive, after), $\alpha = 0.86$ (negative, before), and $\alpha = 0.89$ (negative, after).

Additionally, the following measures were included as part of the coaching program:

Commitment to the goal. Towards the end of each session (except for sessions 1 and 2 in the pilot study), and as part of the coaching session, the CAC asked to what extent the participant was committed to achieving the goal that was being designed during the session. Asking about the commitment of the goal to the clients is a common practice in coaching sessions oriented at goal achievement. Participants rated commitment with their goal on a 7-point scale (1- not at all – 7- very much so).

Goal achievement. At the beginning of the last session, and as part of the coaching session, the CAC asked to what extent the participant had achieved their goal so far. Participants rated goal achievement on a 7-point scale (1- not at all – 7- very much so).

At the end of the study, usability, adoption intention, and satisfaction with the coaching program were evaluated using the following measures:

Usability. A usability questionnaire was developed based on Brooke (1996) to evaluate the usability of the CAC app. The adapted scale contained 6 items (e.g., *The app is easy to use; I would not require the help of technical support to use the app*). Participants rated the items on a 7-point scale (1- strongly disagree – 7- strongly agree). Reliability for the scale was $\alpha = 0.79$.

Adoption intention. A questionnaire to evaluate the willingness to adopt the CAC in real life was developed based on the subscales of perceived usefulness, perceived sociability, and trust from Heerink et al. (2009). The adapted scale contained 5 items (e.g., *I would like to have NORIKA on my phone; I would use NORIKA often if I had it; I think NORIKA would help me achieve my goals if I had it on my phone*). Participants rated the items on a 7-point scale (1 - strongly disagree – 7 - strongly agree). Reliability for the scale was $\alpha = 0.97$.

Satisfaction with the coaching program. A version of the Client Satisfaction Questionnaire (CSQ-8) from Attkisson and Zwick (1982) was adapted to the needs of the study. The adapted scale contained 6 items (e.g., *I am satisfied with the coaching sessions from NORIKA; I would recommend NORIKA to a friend; NORIKA helped me to deal with my goals effectively*). Participants rated the items on a 7-point scale (1- strongly disagree – 7- strongly agree). Reliability for the scale was $\alpha = 0.96$.

Open-ended comments. Participants were encouraged to provide feedback and comments regarding their experience with the coaching sessions and the application.

6. Pilot study

6.1. Participants

Four participants, three females and a male, aged 24–37 ($M = 31.5$), who reported not being diagnosed with any severe psychological disorder and not having used the services of a professional coach before, volunteered for the pilot study. Three participants had completed bachelor's or master's degrees, and one participant had completed a Ph. D. The participants in the pilot study did not receive economic

[1] All Cronbach's α values are calculated using the data from the main study

compensation for their participation.

6.2. Procedure

Upon arrival at the faculty, the participant was led to a quiet room and was invited to sit down in a chair with a table in front. Then, the participant was introduced to the purpose of the study and filled out the consent form. Following this, the participant was introduced to the coaching application. Then, the experimenter played a 10-second demo for each available voice, the female-like and the male-like voices, and asked the participant which of the two they wanted to be their coach. The three female participants chose the female-like voice and the male participant chose the male-like voice. The experimenter loaded the first coaching session using the voice requested by the participant and placed the smartphone on the table, in front of the participant.

Once the session started, the experimenter stayed in a corner of the room and did not intervene in the process. Then, the coaching session took place. The rest of the sessions unfolded in a similar manner.

Before starting the intervention in session 1, the CAC asked the participant to fill out the PGI, the SLS, and the PANAS scales on the smartphone screen. In sessions 3 and 4, when the participant was working towards building their goal, the CAC also collected how committed the participant was to achieve their goal. At the beginning of session 4, the CAC additionally asked the participant to what extent they had achieved their goal so far.

The length of the sessions varied slightly across participants, depending on their pace to complete the different steps and activities, but generally lasted about 12–14 min.

After completing the full coaching program, at the end of the last session, participants answered PGI, SLS, and PANAS on the screen again.

They also completed a series of questions related to usability, adoption intention, and satisfaction with the coaching program, on a separate computer.

Furthermore, the participants were interviewed by one of the experimenters at the end of the study to help identify flaws in the application and the coaching program.

6.3. Results

A paired-sample *t*-test was used to examine modifications in PGI, SLS, and PANAS before and after the intervention.

Participants showed a significant increase in PGI ($p = .013$) and positive affect ($p = .025$) after completing the program compared to before, which support H_1 and H_{3a}. While not significant, there was also a moderate increase in SLS ($p = .124$) and a moderate decrease ($p = .098$) in negative affect, which are in line with H_2 and H_{3b}. The results are reported in Table 2.

Participants also showed a high commitment to their goal, both in session 3 ($M = 5.5$, SD = 0.6) and session 4 ($M = 5.5$, SD = 0.6). Goal achievement at the last session was rated on average as $M = 4.25$ (SD = 1.5).

Usability of the application ($M = 6.53$, SD = 0.45), satisfaction with the coaching program ($M = 6.06$, SD = 0.21), and adoption intention ($M = 5.75$, SD = 0.35) were all rated high by the participants.

Table 2
PGI, SLS, PANAS before and after the coaching program for the pilot study.

	Before		After		t	*p*
	M	*SD*	*M*	*SD*		
PGI	4.5	.35	5.83	.5	−5.28	.013
SLS	4.87	.59	5.62	.14	−2.12	.124
PANAS-P	5.4	.49	5.97	.69	−4.17	.025
PANAS-N	2.2	.82	1.37	.35	2.37	.098

6.4. Modifications in the coaching app after the pilot study

In the course of the interviews, the following aspects emerged. First, some participants pointed out at several aspects of the coaching scripts that required clarification. Second, some participants suggested the possibility of joining Session 1 and 2 in a single session as Session 1 was very introductory and they would have preferred to advance further in the coaching program from the first session. Third, some participants mentioned that they would have liked to have more privacy during the coaching sessions as the presence of the experimenter in the same room inhibited their conversations with the CAC. All these aspects were addressed in the main study. Specifically, the sessions were revised for clarity and restructured into three instead of four sessions. Participants were also given more privacy in the main study (see *7.2. Procedure* for the Main Study). Additionally, we added the measurement of commitment with the goal also in the script for Session 1 of the main study, so that this measure was consistently collected across sessions.

7. Main study

7.1. Participants

Thirty-two participants took part in the study and thirty completed all sessions. The final sample ($n = 30$) included 16 females, 13 males, and 1 non-binary participants. The age range was 20–66 ($M = 40.47$, SD = 9.66). Seven participants had completed studies below university degrees, 19 had bachelor's or master's degrees, and 4 had completed a Ph.D. All participants reported not being diagnosed with any severe psychological disorder and not having used the services of a professional coach before. The participants in the main study received 15 Euros for their participation.

7.2. Procedure

The main study was conducted in the participants' homes, following similar studies (Jeong et al., 2020). The participant was previously requested to have ready a quiet room with a table and a chair where they would not be disturbed during the process and an adjacent room also with no-one else on it where the experimenter could wait during the study. Upon arrival at the participant's home, the experimenter invited the participant to sit down in the chair, introduced the purpose of the study, and requested the participant to fill out the consent form. After that, the experimenter loaded the first coaching session and placed the smartphone on the table, in front of the participant.

Once the session started, the experimenter moved to an adjacent room to offer more privacy to the participant. The experimenter did not intervene in the process unless the participant experienced some technical doubt or difficulty. The subsequent sessions unfolded in a similar manner. Participants were requested to complete sessions 2 and 3 in the same space as session 1.

Before starting with the coaching program in session 1, the CAC asked the participant to fill out the PGI, the SLS, and the PANAS scales on the smartphone screen. Then, the coaching session took place. Towards the end of each session, the CAC collected how committed the participants were to achieve their goal. At the beginning of session 3, the CAC additionally asked the participant to what extent they had achieved their goal so far.

After completing the full coaching program, at the end of the last session, the participants answered PGI, SLS, and PANAS on the screen again. Additionally, the participants completed pen and paper questionnaires related to usability, adoption intention, and satisfaction with the coaching program.

The length of the sessions varied slightly across participants, depending on their pace to complete the different steps and activities, but generally lasted about 15–20 min.

Due to the pandemic situation, additional safety protocols were

undertaken for this study. In particular, both the participant and the experimenter wore face masks and kept a large interpersonal distance at all times. The experimenter wore globes to manipulate the smartphone. The smartphone was placed in a sealed plastic bag. After each session, the bag was replaced by a new one and the smartphone was wiped with rubbing alcohol.

7.3. Results

A paired-sample *t*-test was used to examine modifications in PGI, SLS, and PANAS before and after the intervention.

Participants showed a significant increase in PGI ($p = .001$) and in SLS ($p < .001$) and a significant decrease in negative affect ($p < .001$) after completing the program compared to before, which support H_1, H_2, and H_{3b}. While not significant ($p = .596$), there was also a moderate increase in positive affect, in line with H_{3a}. The results are reported in Table 3.

Participants also showed a high commitment with their goal ($M = 6.17$, SD = 1.01, session 1; $M = 6.03$, SD = 0.99, session 2; $M = 6.4$, SD = 0.85, session 3). Goal achievement at the last session was rated on average as $M = 4.7$ (SD = 1.29).

Usability of the application ($M = 6.14$, SD = 0.77) was rated very high, and satisfaction with the coaching program ($M = 5.12$, SD = 1.29), and adoption intention ($M = 4.33$, SD = 1.84) were rated medium-high by the participants.

8. Overall discussion

We presented a speech-based conversational agent coach that is capable of delivering a basic coaching program for goal achievement based on coaching techniques. We tested its effectiveness in the pilot study and the main study, with a total of 34 participants that completed the full program.

The results of the two studies suggest that the application was able to effectively deliver the coaching program to the participants, engage them with their goals, and contribute to their achievement. While not all measures were significant in both studies, participants reported a consistent increase in PGI, SLS, and positive affect, and a consistent decrease in negative affect, in line with our hypotheses, which suggests that the coaching program positively contributed to overall well-being, life satisfaction, and personal growth. Furthermore, in response to our research questions regarding usability, adoption intention, and satisfaction with the program, the participants from both studies reported a high (pilot) and medium-high (main study) level of satisfaction with the coaching program, as well as a high (pilot) and medium-high (main study) interest to adopt the application in real life. The usability of the application was rated very positively in both studies.

The results of the study also support previously existing literature that has highlighted the benefits of working on goal achievement for psychological well-being (Brunstein, 1993; Spence and Grant, 2007) and further contribute to empirically validating the effectiveness of coaching.

In recent years, an increasing number of smartphone applications to enhance mental health have become available to users. However, these applications are not always supported by empirical evidence that demonstrates their effectiveness (Chittaro and Vianello, 2016; Radovic et al., 2016) and do not always use solid evidence-based techniques, which can be counterproductive for the end-user (Aymerich-Franch and Johnston, 2019). The results of the study are also encouraging towards the possibility of using conversational agents for well-being improvement, which could contribute to universalize access to validated resources for goal achievement and psychological well-being enhancement.

While the capabilities of conversational agents for contributing to well-being improvement might still be limited, especially compared to a human coach or therapist, they might also offer a series of advantages. In particular, there has been a historical reluctance to seek professional help to enhance psychological well-being, which is mostly explained by fear of stigmatization and of being judged, as well as by the elevated costs of visiting a professional (Corrigan, 2004; Sareen et al., 2007; Wang et al., 2005). Conversational agents might be free from the stigma that has traditionally been associated with receiving psychological help from a professional. Users might perceive these agents as unable to judge them for being a machine and not a human being and thus find it easier to share personal information with them compared to a human coach. In this regard, some participants mentioned in the open-ended comments that: "I would rather have NORIKA as my coach because sometimes you can feel shy to talk to a real person to share your stuff" or "I think someone that has difficulty to talk to a real human might find it easier to talk to NORIKA because you don't feel judged". Furthermore, conversational agents for well-being could also offer much more flexibility than a human professional as they could be used anywhere, any time. Also, people that generally prefer using self-help methods might benefit from a significantly more advanced technology than the current existing platforms for self-help. In addition, an agent of these characteristics would be much more affordable than a professional coach, or even free, for the end-user.

That said, while most participants accepted the CAC as an effective virtual coach, several also expressed a preference for a human coach over a virtual one in the open-ended comments, were they to choose between both (e.g., "I think I'd feel better with a human than with a machine, I missed the human warmth").

The study presents some limitations which should be taken into consideration. First, we did not include a control condition in the main study as initially planned. The main study was moved to the participants' homes as a result of the pandemic situation. Since the surrounding environment of the participants remained constant only across sessions but not across subjects a mixed design was discarded. However, the lack of a control group leaves an open door to the question of whether or not the observed effects are really being caused by the intervention. Also, the before-after design may have induced the participants to feel that they had to report making progress and enhanced well-being to content the researchers after completing the study. These limitations are also reported in previous studies that have used a similar design (Grant, 2003; Jeong et al., 2020).

Additionally, there are some aspects of the application that can be improved. Specifically, some participants mentioned that the "voice sounded robotic" and that the CAC "needed to be more interactive". Others also requested that ideally they "would have liked more sessions". Future work will involve working on the capacity of the CAC to become more interactive and on improving the naturality of its voice.

Ultimately, the results of the study provide an empirically-validated program for coaching interventions delivered by autonomous agents, add to previous literature that highlights the benefits of coaching for the psychological well-being (Grant, 2003, 2009; Green et al., 2006; Spence and Grant, 2007), and contribute to further explore the applied uses of emerging media for well-being enhancement (Reinecke and Eden, 2017).

Table 3
PGI, SLS, PANAS before and after the coaching program for the main study.

	Before		After		t	*p*
	M	*SD*	*M*	*SD*		
PGI	4.74	1.00	5.47	1.09	−3.84	.001
SLS	4.97	.85	5.65	.90	−4.99	<0.001
PANAS-P	5.92	.70	6.02	.86	−0.54	.596
PANAS-N	2.23	1.05	1.54	.83	4.31	<0.001

Credit author statement

Laura Aymerich-Franch: conceptualization, design and content, methodology, validation, formal analysis, investigation, resources,

writing - original draft, writing - review & editing, visualization, supervision, funding acquisition

Iliana Ferrer: investigation, data curation, writing - review & editing

Funding

The project has received funding from the Ramón y Cajal Fellowship Program (ref. RYC-2016–19770) awarded to the first author, funded by Agencia Estatal de Investigación, Ministerio de Ciencia, Innovación y Universidades, and the European Social Fund.

CRediT authorship contribution statement

Laura Aymerich-Franch: Conceptualization, Methodology, Validation, Formal analysis, Investigation, Resources, Writing – original draft, Writing – review & editing, Visualization, Supervision, Funding acquisition. **Iliana Ferrer:** Investigation, Data curation, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data Availability Statement

The data supporting the findings of this study are available within the article and its supplementary materials (SI – dataset).

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.ijhcs.2021.102745.

References

Albaina, I.M., Visser, T., Van Der Mast, C.A.P.O., & Vastenburg, M.H. (2009). Flowie: a persuasive virtual coach to motivate elderly individuals to walk. 2009 3rd International Conference on Pervasive Computing Technologies for Healthcare - Pervasive Health 2009, PCTHealth 2009. https://doi.org/10.4108/ICST.PERVASIVEHEALTH2009.5949.

Alberts, H., & Poole, L. (2019). Goal Visualization. Positive Psychology Practitioners Toolkit. www.positivepsychologyprogram.com.

Attkisson, C.C., Zwick, R., 1982. The client satisfaction questionnaire. Eval. Program Plann. 5 (3), 233–237. https://doi.org/10.1016/0149-7189(82)90074-X.

Aymerich-Franch, L., & Bailenson, J. (2014). The use of doppelgangers in virtual reality to treat public speaking anxiety: a gender comparison. Proceedings of the International Society for Presence Research Annual Conference. March, 17-19, Vienna, Austria. – Top Three Paper Award.

Aymerich-Franch, L., & Johnston, T. (2019). Intervention framework for a robot coach to improve psychological well-being. International Communication Association Conference (ICA).

Ayres, J., Hopf, T., 1993. Coping with speech anxiety. The Communication and Information Science Series.

Brooke, J., 1996. SUS - a quick and dirty usability scale. Usability Eval. Industry 189 (194), 4–7. https://doi.org/10.1002/hbm.20701.

Brunstein, J.C., 1993. Personal goals and subjective well-being: a longitudinal study. J. Pers. Soc. Psychol. https://doi.org/10.1037/0022-3514.65.5.1061.

Chittaro, L., Vianello, A., 2016. Mobile mindfulness and user's worry: a qualitative study of using a smartphone app for distancing from negative thoughts. Interact. Comput. https://doi.org/10.1093/iwc/iwv044.

Corrigan, P., 2004. How stigma interferes with mental health care. Am. Psychol. 59 (7), 614–625. https://doi.org/10.1037/0003-066X.59.7.614.

Cox, E., Bachkirova, T., Clutterbuck, D., 2014. The complete handbook of coaching. The Complete Handbook of Coaching, 2nd Ed.

Diederich, S., Brendel, A.B., Kolbe, L.M., 2019. On conversational agents in information systems research: analyzing the past to guide future work. In: Proceedings of the International Conference on Wirtschaftsinformatik.

Diener, E., Emmons, R.A., Larsen, R.J., Griffin, S., 1985. The satisfaction with life scale. J. Pers. Assess. 49 (1), 71–75. https://doi.org/10.1207/s15327752jpa4901_13.

Doran, G.T., 1981. There's a S.M.A.R.T. way to write management's goals and objectives. Manage. Rev.

Frisch, M.B., 2006. Quality of Life therapy: Applying a Life Satisfaction Approach to Positive Psychology and Cognitive Therapy. Wiley, Hoboken.

Frisch, M.B., 2016. Quality of life therapy. The Wiley Handbook of Positive Clinical Psychology. https://doi.org/10.1002/9781118468197.ch27.

Grant, A.M., 2003. The impact of life coaching on goal attainment, metacognition and mental health. Soc. Behav. Personal. 31 (3), 253–263. https://doi.org/10.2224/sbp.2003.31.3.253.

Grant, A.M., Cavanagh, M.J., 2010. Life coaching. In: Cox, E., Bachkirova, T., Clutterbuck, D. (Eds.), The Complete Handbook of Coaching.

Grant, A.M., Curtayne, L., Burton, G., 2009. Executive coaching enhances goal attainment, resilience and workplace well-being: a randomised controlled study. J. Positive Psychol. https://doi.org/10.1080/17439760902992456.

Green, S., Oades, L.G., Grant, A.M., 2006. Cognitive-behavioral, solution-focused life coaching: enhancing goal striving, well-being, and hope. J. Positive Psychol. T 1 (3), 142–149. https://doi.org/10.1080/17439760600619849.

Heerink, M., Kröse, B., Evers, V., Wielinga, B., 2009. Measuring acceptance of an assistive social robot: a suggested toolkit. In: Proceedings - IEEE International Workshop on Robot and Human Interactive Communication, pp. 528–533. https://doi.org/10.1109/ROMAN.2009.5326320.

Holzinger, A., Searle, G., Kleinberger, T., Seffah, A., Javahery, H., 2008. Investigating usability metrics for the design and development of applications for the elderly. In: International Conference on Computers for Handicapped Persons. Springer, Berlin, Heidelberg, pp. 98–105.

Hoy, M.B., 2018. Alexa, Siri, cortana, and more: an introduction to voice assistants. Med. Ref. Serv. Q. https://doi.org/10.1080/02763869.2018.1404391.

Hudlicka, E., 2013. Virtual training and coaching of health behavior: example from mindfulness meditation training. Patient Educ. Counsel. https://doi.org/10.1016/j.pec.2013.05.007.

Ives, Y., 2008. What is "Coaching"? An exploration of conflicting paradigms. Int. J. Evid. Based Coach. Mentoring.

Jeong, S., Alghowinem, S., Aymerich-Franch, L., Arias, K., Lapedriza, A., Picard, R., Park, H.W., & Breazeal, C. (2020). A robotic positive psychology coach to improve college students' wellbeing. 29th IEEE International Conference on Robot and Human Interactive Communication, RO-MAN 2020. https://doi.org/10.1109/RO-MAN47096.2020.9223588.

Kamphorst, B.A. (2017). E-coaching systems: what they are, and what they aren't. Personal and Ubiquitous Computing. https://doi.org/10.1007/s00779-017-1020-6.

Kreps, G.L., Neuhauser, L., 2010. Editors' introduction, Ehealth and health promotion. J. Comput.-Mediated Commun. https://doi.org/10.1111/j.1083-6101.2010.01526.x.

Ly, K.H., Ly, A.M., Andersson, G., 2017. A fully automated conversational agent for promoting mental well-being: a pilot RCT using mixed methods. Internet Intervent. https://doi.org/10.1016/j.invent.2017.10.002.

Magyar-Moe, J.L., 2009. Therapist's guide to positive psychological interventions. Therapist's Guide to Positive Psychological Interventions. https://doi.org/10.1016/S1873-0450(09)X0003-3.

Merdivan, E., Singh, D., Hanke, S., Kropf, J., Holzinger, A., Geist, M., 2020. Human annotated dialogues dataset for natural conversational agents. Appl. Sci. 10 (3), 762.

Pew Research Center (2017). "Nearly half of Americans use digital voice assistants, mostly on their smartphones". Available at: http://www.pewresearch.org/fact-tank/2017/12/12/nearly-half-of-americans-use-digital-voice-assistants-mostly-on-their-smartphones [last access: February 21st, 2021].

Radovic, A., Vona, P.L., Santostefano, A.M., Ciaravino, S., Miller, E., Stein, B.D., 2016. Smartphone applications for mental health. Cyberpsychol., Behav. Soc. Networking. https://doi.org/10.1089/cyber.2015.0619.

Reinecke, L., Eden, A., 2017. Media use and well-being: an introduction to the special issue. J. Media Psychol. https://doi.org/10.1027/1864-1105/a000227.

Riva, G., Baños, R.M., Botella, C., Wiederhold, B.K., Gaggioli, A., 2012. Positive technology: using interactive technologies to promote positive functioning. Cyberpsychol., Behav. Soc. Network. 15 (2), 69–77. https://doi.org/10.1089/cyber.2011.0139.

Robitschek, C., 1998. Personal growth initiative: the construct and its measure. Measur. Evaluat. Counsel. Development 30 (4), 183–198. https://doi.org/10.1023/A.

Ruvolo, A.P., Markus, H.R., 1992. Possible selves and performance: the power of self-relevant imagery. Soc. Cogn. https://doi.org/10.1521/soco.1992.10.1.95.

Sareen, J., Jagdeo, A., Cox, B.J., Clara, I., ten Have, M., Belik, S.-.L., de Graaf, R., Stein, M.B., 2007. Perceived barriers to mental health service utilization in the United States, Ontario, and the Netherlands. Psychiatr. Serv. 58 (3), 357–364. https://doi.org/10.1176/ps.2007.58.3.357.

Spence, G.B., Grant, A.M., 2007. Professional and peer life coaching and the enhancement of goal striving and well-being: an exploratory study. J. Positive Psychol. https://doi.org/10.1080/17439760701228896.

Stoltzfus, T., 2008. Coaching questions. A Coach's Guide To Powerful Asking Skills, p. Coach22.

ter Stal, S., Kramer, L.L., Tabak, M., op den Akker, H., Hermens, H., 2020. Design features of embodied conversational agents in eHealth: a literature review. Int. J. Hum. Comput. Stud. 138, 102409.

Wang, P.S., Berglund, P., Olfson, M., Pincus, H.A., Wells, K.B., Kessler, R.C., 2005. Failure and delay in initial treatment contact after first onset of mental disorders in the National Comorbidity Survey Replication. Arch. Gen. Psychiatry 62 (6), 603–613. https://doi.org/10.1001/archpsyc.62.6.603.

Watson, A., Bickmore, T., Cange, A., Kulshreshtha, A., Kvedar, J., 2012. An internet-based virtual coach to promote physical activity adherence in overweight adults: randomized controlled trial. J. Med. Internet Res. https://doi.org/10.2196/jmir.1629.

Watson, D., Clark, L.A., Tellegen, A., 1988. Development and validation of brief measures of positive and negative affect. PANAS Scales 54, 1063–1070. https://doi.org/10.1037/0022-3514.54.6.1063.

Whitmore, J., 2010. Coaching for performance - growing human potential and purpose. NHRD Network J. https://doi.org/10.1177/0974173920100216.

Whitten, P., Steinfield, C., Hellmich, S., 2001. Ehealth: market potential and business strategies. J. Comput.-Med. Commun. https://doi.org/10.1111/j.1083-6101.2001.tb00129.x.