Robots as Mental Well-being Coaches: Design and Ethical Recommendations

MINJA AXELSSON, AFAR Lab (Affective Intelligence and Robotics Laboratory); Department of Computer Science and Technology; University of Cambridge, United Kingdom
MICOL SPITALE*, Politecnico di Milano, Italy
HATICE GUNES, AFAR Lab (Affective Intelligence and Robotics Laboratory); Department of Computer Science and Technology; University of Cambridge, United Kingdom

The last decade has shown a growing interest in robots as well-being coaches. However, insightful guidelines for the design of robots as coaches to promote mental well-being have not yet been proposed. This paper details design and ethical recommendations based on a qualitative analysis drawing on a grounded theory approach, which was conducted with a three-step iterative design process which included user-centered design studies involving robotic well-being coaches, namely: (1) a user-centred design study conducted with 11 participants consisting of both prospective users who had participated in a Brief Solution-Focused Practice study with a human coach, as well as coaches of different disciplines, (2) semi-structured individual interview data gathered from 20 participants attending a Positive Psychology intervention study with the robotic well-being coach Pepper, and (3) a user-centred design study conducted with 3 participants of the Positive Psychology study as well as 2 relevant well-being coaches. After conducting a thematic analysis and a qualitative analysis, we collated the data gathered into convergent and divergent themes, and we distilled from those results a set of design guidelines and ethical considerations. Our findings can inform researchers and roboticists on the key aspects to take into account when designing robotic mental well-being coaches.

CCS Concepts: • **Computer systems organization** → **Robotics**; • **Human-centered computing** → *HCI design and evaluation methods*; User studies; **User centered design**; *HCI theory, concepts and models*; • **Computing methodologies** → *Artificial intelligence.*

Additional Key Words and Phrases: Human-Robot Interaction, Social Robots, Robot Design, Design Guidelines, Ethical Considerations

*This work was undertaken and finalised while Micol Spitale was a postdoctoral researcher at the University of Cambridge.

Authors' addresses: Minja Axelsson, minja.axelsson@cl.cam.ac.uk, AFAR Lab (Affective Intelligence and Robotics Laboratory); Department of Computer Science and Technology; University of Cambridge, Cambridge, United Kingdom; Micol Spitale, micol.spitale@polimi.it, Politecnico di Milano, Milan, Italy; Hatice Gunes, hatice.gunes@cl.cam.ac.uk, AFAR Lab (Affective Intelligence and Robotics Laboratory); Department of Computer Science and Technology; University of Cambridge, Cambridge, United Kingdom.

1 INTRODUCTION

The prevalence of mental health conditions has grown significantly in the last decade – with an increase of 13% in people diagnosed with mental health disorders [81] – and it has been even more exacerbated by the COVID-19 pandemic. This increase is due to several risk factors, such as the use of social media, the inequality and social disadvantage, discrimination and social exclusion, traumatic experiences, differences in physical health etc. [43, 89]. These conditions can have a negative impact on people's everyday life, affecting for example their academic or professional achievements, relations with family and friends, and their community involvement. Despite these numbers, less than 2% of worldwide government health spending is allocated to mental health even if more would be needed to meet the request of mental health services. The limitations in accessibility of mental well-being practices, the lack of personnel and resources to support people in this regard, and the advent of mental health digital applications (e.g., mobile apps) — widely used especially during the COVID-19 pandemic [3, 15] — have contributed to the increase of research interest in the use of robots as coaches for well-being [92].

To respond to those requests, many tech companies offered affordable and accessible mental well-being service, such as apps (e.g., HeadSpace, Calm) that have been shown to have a great potential to help the promotion of people's mental well-being [11, 91]. Despite this, recent studies [63, 97] indicated that participants struggled to find time to use a meditation app, and that a high dropout rate and low app engagement represent well-known hurdles [66]. On the contrary, past works in Human-Robot Interaction (HRI) have showed that robots can help in promoting mental health application in the long term, such as for mindfulness exercises [17] or positive psychology practices [91], as well as supporting physical well-being, such as in rehabilitative physical therapy [104], even if they may be expensive with respect to app-based solutions. Previous works have investigated socially assistive robots from the perspective of well-being, care, and justice, and has called for the formulation of practical principles for SAR design [16]. As explorations of robots as mental well-being coaches are still quite novel, cohesive guidelines for their design do not exist. Guidelines have been previously presented to share knowledge on how to design robots for specific applications, such as for communication therapy for children with autism [10], and for encouraging children's creativity [4]. In this paper, we aim to present such design guidelines for robotic mental well-being coaches, to further inform the field on how to design them.

To this end, we have followed an iterative design process by conducting three different user-centred design studies exploring a robotic mental well-being coach. In the first study, we involved 11 participants – including University's students –who had previous experience with well-being practices. We collected their feedback for a hypothetical robot (i.e., participants were not asked to think about a specific, pre-existing robot), gathering a broad range of user perspectives and needs on different types of well-being practices a robotic coach could conduct. In the second study, building upon the first study results, we conducted interviews with 20 participants – recruited among the University's students – who had just interacted with a robotic well-being coach conducting positive psychology sessions. In the third study, 5 participants (both prospective users and coaches) were invited to have a larger group discussion based on their experience in the previous study, and then to explicitly edit robotic coach dialogue.

We analyzed the data collected with Thematic Analyis (TA) [20], and then we conducted a qualitative analysis inspired by the Qualitative Meta-Analysis method [96] of the three studies. The results included what advantages (e.g., improved accessibility and lack of judgement) and disadvantages (e.g., privacy concerns and lack of humanness) prospective users expect to have in a potential robotic mental well-being coach. Additionally, participants detailed features and capabilities they would expect to see in such a robot. Then, we operationalized those results as

recommendations and guidelines, in order to contribute to the future design of such robots. The major contributions of this paper are the following:

- we report results from three user-centered design studies highlighting the design features and capabilities that a robot should have to be used as a well-being coach;
- based on the lessons learned from the studies, we identified a set of recommendations in terms of design guidelines and ethical considerations to inform researchers and roboticists about the key aspects to take into account during the design process of a robotic coach for mental well-being.

The rest of the paper is organized as follows. Section 2 summarizes the background and related works that informed this research, Section 3 describes the methodology for our research, Section 4.1, 4.2 and 4.3 describe respectively Study 1, Study 2 and Study 3. Section 5 reports the qualitative analysis from the results of the previous three studies, Section 6 reports the design guidelines obtained from the studies, and discusses ethical consideration, and Section 7 concludes the paper and proposes directions for future research.

2 BACKGROUND AND RELATED WORK

In this section, we review relevant literature on Well-Being Coaching (WBC) and other mental well-being practices, previous studies examining well-being applications in Human-Robot Interaction (HRI), as well as the use of User-Centred Design (UCD) and Qualitative Research in HRI for well-being.

2.1 Well-being coaching and other mental well-being practices

Well-being coaching is a non-clinical practice which focuses on improving the coachee's mental well-being, goal-striving and hope [48]. Coaching (cf. psychological therapy aimed to address mental illness) emphasizes the present, as well as how the coachee may flourish in the future, and maximize their fulfilment in life and work [52]. The usual structure of coaching in a one-to-one session is to clarify the coachee's issues, establish clear, measurable and attainable goals, develop action plans, and review progress in the next session [47, 48]. Issues addressed by coaching can range from lifestyle (e.g. smoking cessation), to working life (e.g. learning a new skill), to personal relationships (e.g. addressing an issue with a partner). The coach's goal is to facilitate and activate the coachee's thinking and progress toward the goal. Coaching can emphasize different psychological practices, e.g. cognitive behavioural (focus on the relationship between thoughts, feelings, and actions) and brief-solution focused practice (helping the coachee focus on how their problem is already being addressed in small ways) [48], or positive psychology (paying greater attention to the positive aspects of one's life) [88]. Another style of coaching is Life Coaching, where goals can include dealing with stress, creating more meaningful relationships, and creating a more fulfilling and purposeful life in general. Life coaches can use varying tools such as drawing, writing, body-awareness exercises or visualization [46].

Other forms of well-being practices that serve to improve the practitioner's mental well-being are mindfulness and meditation, which can improve the practitioner's well-being and reduce negative emotions [36]. Mindfulness training is often integrated into a meditation practice, and such meditation has shown to e.g. decrease stress and anxiety in college students [13]. Mindfulness has experienced a rise in popularity recently, with it being applied to workplace environments [58] as well as seeing new digital applications to conduct mindfulness training [77]. An important difference between mindfulness training and well-being coaching, when considering from a digital implementation perspective, is that the first can be conceptualized as instructing the participant in a particular skill, while the second is a conversation-based practice that trains people to focus on small

things in their lives. This distinction may lead to different design recommendations for robotic well-being coaches operating in this area. It should be noted that both forms of well-being practices have already been recreated and commercially available in mobile application form: the mindfulness meditation application Headspace [54] has been shown to improve depressive symptoms and improve resilience [41], and the application Woebot was applied to deliver cognitive-behavioural therapy to young adults and reduced depressive symptoms [40].

2.2 Mental Well-being in HRI

While robots have been explored for clinical interventions (such as communication therapy for children with autism spectrum disorders (ASD) [87] and social commitment robots for elderly people with dementia [73]), applications focusing on mental well-being for non-clinical populations of different ages are very few.

For example, Bodala et al. [17] examined the participants' perceptions of a robotic vs. a human mindfulness coach in a longitudinal study, finding that while the human coach was rated significantly higher than the robotic coach, both received positive feedback. The study targeted a non-clinical population and excluded participants with high anxiety or depression levels prior to the study. Additionally, participants' *Neuroticism* (i.e., nervousness, moodiness, and temperamentality) and *Conscientiousness* (i.e., organization, thoroughness, and reliability) personality traits (as defined in the OCEAN model of personality [45, 80]) were found to influence their perceptions of the robot. Also, Jeong et al. [57] conducted a longitudinal study with the robot Jibo enabling positive psychology interventions for students, showing improvements in participants' well-being, mood, and readiness to change. Findings from this study also confirmed that participants' personality traits of neuroticism and conscientiousness influenced the interventions' efficacy. This study also targeted a non-clinical population, although no selection method is mentioned. Robots have been also used to disclose feeling and thoughts. For instance, Akiyoshi et al. [2] explored how to make people's moods more positive through conversations about their problems with a robot, as compared before and after the conversation with the robot. Their findings showed that people who interacted with the robot self-disclosed more and experienced less anger than those who did not use the robot. Duan et al. [34] ran an empirical study to compare self-disclosing in a diary journal or to a social robot after negative mood induction, targeting a population of people with depression. Their results showed that people who felt strongly negative after the negative induction benefited the most from talking to the robot, rather than writing down their feelings in the journal. Additionally, a preliminary study for a robot displaying mood data, applying methods of community-based discussions with students, was conducted to improve the data visualization methods of such a robot [62]. Arts-based theatre methods were used in a pilot study with a Nao robot to address well-being with older adults living in a residential care setting [38]. The 3-session study found that depression and loneliness scores decreased in the study population during the study, which involved interacting with a robot. A recent study also explored how a robot can aid in the evaluation of children's mental well-being [1]. These results demonstrated that the mental well-being evaluation with the robot seems to be the most suitable in identifying well-being related anomalies as compared with the self-report and the parent-report standard tests.

2.3 User-centred methods in HRI for well-being

User-centred HRI aims to incorporate users' perspectives and needs into the design of robotic applications [64].

Participatory Design (PD) refers to actively involving stakeholders (e.g. prospective users and relevant subject domain experts such as well-being coaches) in the design process as active co-designers [90]. User-centred design has been previously used to design applications aimed at mental

well-being in HRI, mainly in clinical contexts. Some authors also noted creating design guidelines or recommendations for robots in their respective applications at the end of the design process. Lee et al. [67] employed PD to design robots for various social contexts together with older adults with depression, taking into account the participants' issues and concerns in a socially responsible way. Winkle et al. [104] used PD to design robots for rehabilitative therapy together with therapists in focus groups, and presented design implications for robots and best practices based on therapists' knowledge. Axelsson et al. [10] used PD to design a robot to teach sign language to children with autism. The authors presented design guidelines for future robots for children with ASD. Moharana et al. [72] employed PD to design robots for dementia caregiving together with dementia caregiver support groups, broadening the context of dementia robot design to include informal family caregivers in the process. The authors introduced design guidelines for such robots, and elucidated how robot attributes should change according to the stage of dementia of the caregivee.

To the best of our knowledge, none of the past works have defined a set of design and ethical guidelines for robotic mental well-being coaches that can inform the field on how to design and develop them. This paper leveraged on the results of three user-centered studies involving human mental well-being coaches and coachees to distil this set of guidelines as described in the following sections.

3 METHODOLOGY

This section introduces the concepts of Thematic Analysis (TA), which we use to collate and analyze the data from the three studies conducted (described in detailed respectively in Section 4.1, 4.2 and 4.3), as well as a qualitative analysis inspired by the Qualitative Meta-Analysis, which we employ to present a cohesive whole out of these three studies, as well as what we base our design guidelines and recommendations on.

3.1 Thematic Analysis

Our three studies use Thematic Analysis (TA) as a method to analyze qualitative data collected from interviews and workshops. This method has been previously used to examine robotic well-being applications [21, 72, 105]. For example, Moharana et al. [72] used grounded theory to organize data collected from interactions with family caregivers into 16 major themes encountered in dementia care. The researchers used the themes to create scenarios of dementia care, which were then used for robot design. In each of the three presented studies, we employ the 6-step method exemplified by Braun and Clarke [20]. The 6-steps consist of: 1) familiarizing ourselves with the data (i.e., transcribing the data, reading them, and noting some initial ideas), 2) creating initial codes (i.e., identifying the codes within the dataset and collating data to the corresponding code), 3) searching for themes (i.e., collating codes into themes collecting all data under the relevant theme), 4) reviewing the themes (i.e., checking if the themes identified work also in relation with the codes), 5) defining and naming the themes (i.e., generating specific definitions and names for each theme, consistently with the story of the whole dataset collected), and 6) finally creating a report (e.g., extrapolating examples for each theme). We applied a grounded theory approach — defined as the analysis of generating a reasonable theory of the phenomena that is grounded in the data [71] — where the themes in all three studies were selected on the basis of what we found in the data.

3.2 Qualitative Analysis

We apply a grounded-theory qualitative analysis inspired by Qualitative Meta-Analysis on the results of these three studies, in order to synthesize a summary, and to examine themes in common, as well as themes that differ and why these differences may arise [95, 96]. We conduct this multi-study analysis to provide a "more comprehensive description" [95] of how robotic well-being

coaches should be designed taking into account users' and professional coaches' perspectives. In our analysis, we take into account the different context of each study, i.e., they were conducted at different phases of the iterative design process. This focus on the analysis of consecutive and iterative design phases makes it distinct from traditional Qualitative Meta-Analysis. To conduct this qualitative analysis, we examined the results of the TA conducted on all three studies, how themes were repeated through each study (congruence), and how the results of the studies differed from each other (divergence). We analyze how these convergences and divergences may have resulted due to either differences in opinion, or differences in the study structures. We do this in order to both highlight the findings of the different studies, as well as how these findings were influenced by the different context in different phases of the iterative design process. Based on this qualitative analysis, we propose design guidelines for the creation of future robotic well-being coaches. We argue that our three studies and the resulting qualitative analysis are sufficient to propose such guidelines, as we observe *data saturation*, i.e. the point where no new information are observed within the data [49]. Our original three studies address the design problem of robotic mental well-being coaches from different perspectives and with different participants, and we see similar themes repeated throughout all studies. Additionally, we pay attention to the differences in these studies, as these can present important contextual differences in specific mental well-being applications or participant populations.

4 ITERATIVE DESIGN PROCESS: THREE USER-CENTRED STUDIES

In order to design the robot coaches to promote mental well-being, we followed a three-step iterative design process consisted of three user-centred studies. Firstly, we held an exploratory user-centred design study (described in Section 4.1) in which both well-being coaches and university students' prospective users of mental well-being practice who had no to very little experience with robots participated. Our results suggested that the robot's ability to adapt and recognize users' emotions were crucial for the success of the practice delivered by a robotic coach. This motivated our second user-centred study (described in Section 4.2), in which participants (again University of Cambridge's students) interacted with a robotic coach featuring adaptive capacities to deliver positive psychology exercises. The collaboration with one of the coaches in Study 1, who had expertise in positive psychology, motivated us to focus on this particular coaching style. We then concluded our design process by conducting another user-centred design study which included the participants of Study 2 – who had interacted with the robotic coach – as well as the well-being coaches (see Section 4.3) and focused on the evaluation of various applications of robotic coaches to promote mental well-being.

4.1 Study 1

The first study aims to collect and examine the expectations and perceptions of prospective well-being users and well-being coaches. This study was originally discussed in [7]. Here, we summarize the study, and present aspects related to well-being practices that were out of scope for the prior publication.

4.1.1 Participants. We involved 11 participants, 8 of them were interested in well-being practices with a robotic coach (as reported by themselves when asked by the researchers), and 3 of them (2 females and 1 male) were well-being coaching professionals, specialized in Solution-Focused Practice (SFP), Mindfulness, and Life Coaching respectively. We conducted group discussions with the 8 prospective users, and semi-structured interviews with all participants, including the 3 well-being coaches. The prospective users had taken part in either SFP sessions with a human coach (6 participants, details of study and practice reported in [7, 24]), or Mindfulness sessions with a human

coach (2 participants, details of study and practice reported in [17]). In the SFP and Mindfulness studies, participants interacted with a coach conducting Mindfulness sessions (5 sessions, in a group) or SFP sessions (4 sessions, individually). Participants in both studies were recruited from the University of Cambridge student and postdoctoral researcher population. Further demographic details of the participants were not collected in this study.

4.1.2 Setting. Due to COVID-19 restrictions, we performed both the group discussion with the prospective users and the semi-structured interviews online (via MS Teams). To facilitate the group discussions in the user-centred session online, we asked participants to use the online tool *Miro*[1] (an example of working with Miro can be seen in Figure 1).

Items	**Duration**
Pre-discussion survey (in writing)	5 min
Introduction	3 min
Warm-up discussion about well-being practices	10 min
Introduction to social robots and demo videos	7 min
Ideating robotic well-being coach	15 min
Discussion on robotic well-being coach features and capabilities	20 min
Conclusion	2 min
Post-discussion survey (in writing)	5 min

Table 1. Structure of the group discussions. Individual 30-45 minute interviews performed beforehand. Originally presented in [7].

4.1.3 Protocol. This study focused on the prospective users' and coaches' experiences when participating in and coaching well-being practices, and how technology has helped or hindered these tasks. It consisted of group discussions, semi-structured interviews, and a canvas filling (optional) [7].

We first conducted 8 one-to-one semi-structured interviews with the prospective users for 30-45 minutes, to ask them about their experiences with well-being practices, well-being practices involving technology, and potential improvements and challenges they had encountered (see Appendix B for an example interview structure). Additionally, we interviewed the 3 well-being coaches about their experiences with different coaching practices, how they had previously used technology, and how they could see a robot conducting coaching (see Appendix C for an example interview structure).

After the interviews, we organized group discussions with the same 8 prospective users we had interviewed. The discussion topics were informed by the interviews conducted with the prospective users and coaches. We divided the group discussion participants into two groups (4 participants each), and each group discussion followed the same structure (see Table 1). Each group discussed their previous experiences with well-being practices, how they could see a robot coaching well-being practices, and what capabilities and features such a robotic coach should have.

The full list of topics discussed during the group discussion was as follows: (1) Well-being practices (e.g., Describe with one word your experience participating in the SFP / Mindfulness study), (2) Aim of the group discussion (e.g., Whether and how social robots can used as coaches for well-being), (3) Ideation on a robotic well-being coach (e.g., How do you think social robots could

[1]https://miro.com

provide or support well-being practices, or work as a well-being coach?), (4) What should the robot be like? (e.g., What do you want the robot to do / say?). Please find more detail of the study in [7].

Three of the prospective users as well as two well-being coaches also participated in an optional part of the study, where they filled in a subset of the Social Robot Co-Design Canvases [6, 9], to further detail design considerations in a robotic well-being coach.

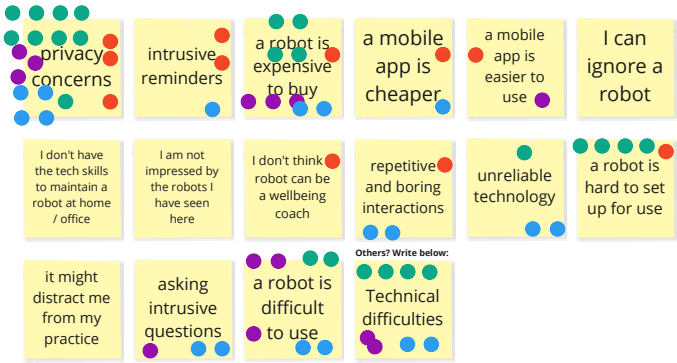

Fig. 1. Results of participant votes using the Miro board to indicate what would stop them from using the robotic well-being coach. Originally published in [7].

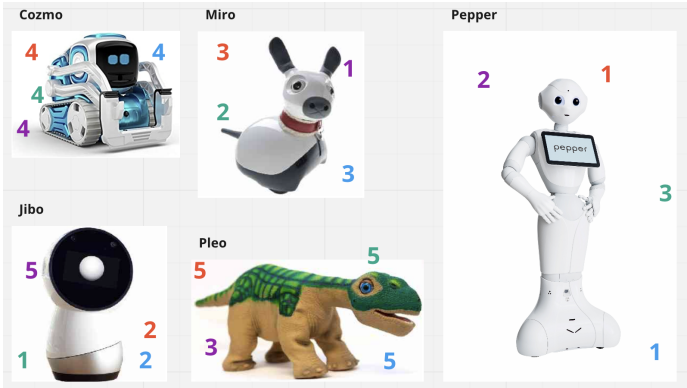

Fig. 2. Results of participants' ranking robot embodiment for well-being coaching (discussion group of four participants, each having their own colour). While Pepper was preferred by this group, Miro and Jibo were also preferred by some, indicating that preferences vary across individuals. Originally published in [7].

4.1.4 Data analysis. A Thematic Analysis (TA) was conducted, following the procedure described in Section 3.1, based on data transcribed from interviews (both with prospective users and coaches) and group discussions, as well as data collected from the Miro board and the Social Robot Co-Design Canvases.

4.1.5 Summary of Main Findings. The TA resulted in 9 major themes (as seen in Figure 3). This section summarised the main findings of the study. Quotes on the topics summarized in this paper as well as images from the group discussions and more detailed and extensive description of the results can be found in [7]. In the following sections, we refer to the prospective users with the

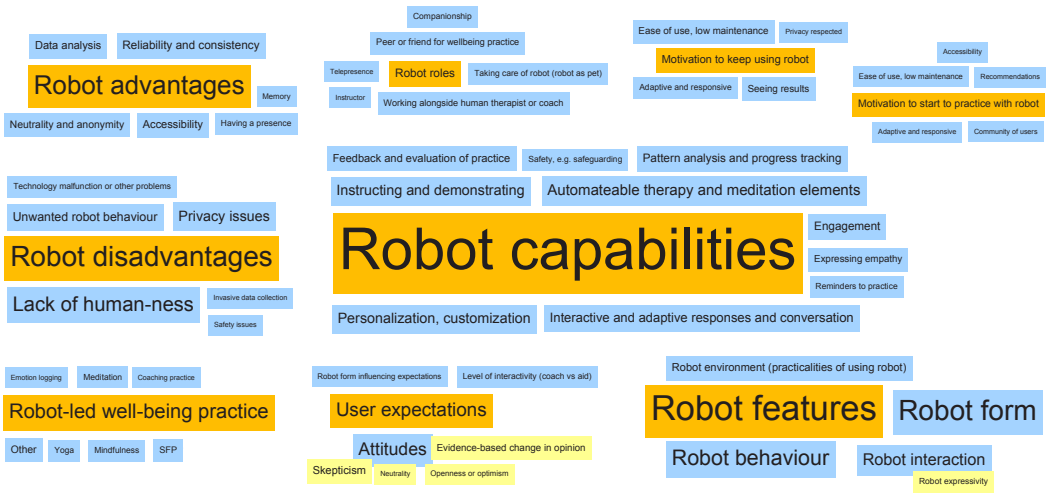

Fig. 3. Study 1: Themes defined in the TA are presented in orange, while codes related to these themes are presented in blue, and sub-codes are presented in yellow (best viewed in colour), from [7].

term "participants" while we specify which quote has been told by the coaches directly naming them.

Well-being practices. Both prospective users and coaches were asked to think about what *well-being practices* they could see a robot leading, bringing up *mindfulness, meditation, yoga* and *SFP*. Participants detailed what kind of *role* the robot should take during such practices (e.g. a *instructor*, a *peer or friend* for the practice, or a *supplement for instructions from a human coach*). Participants perceived that the robot should have differing roles depending on the practice the robot would be conducting, e.g., for meditation the robot could act as more of a peer, and for SFP more as an instructor. The practices discussed were drawn mainly from the participants' own previous experiences with well-being practices, either from their own practice or from the studies they had participated in. Comments made by the three professional coaches about how a robotic coach might conduct the practices they were specialized in can be seen in Table 2, e.g., examining the specific types of SFP techniques (The Miracle Question) and Meditations (Mindfulness of Breathing and Loving Kindness) they could see the robot having an advantage in.

Robot capabilities. A major focus during group discussion and coach interviews was on *robot capabilities* during well-being coaching, i.e., what capabilities participants thought that the robot should display during its interactions with coachees, in order for them to benefit from the interaction. Participants thought that a RC could *engage participants in the practice* by providing a *sense of presence*, or by *doing the practice together with them*. They noted that the physical presence of the robot could help them better focus on the practice (in comparison to e.g. a mobile app, which has no presence). Participants also noted that they would like the robot to *express empathy* by *reacting to how they feel* and to *praise* them in suitable situations. This would help them feel that the robot was acknowledging their emotions. Participants also hoped that the robot would display empathy by changing its *expressions*, however the SFP coach remarked that it would be important to have the correct level of expressivity in order to mitigate the *uncanny valley* effect that could arise due to incongruent responses.

Robot *responsiveness* in terms of *interactive and adaptive responses* was also discussed. Participants said they would like the robot to *summarize their thoughts* to give a sense of *acknowledgement*. A participant noted that such a sense could also be achieved through *expressive behaviours and movement* instead of tailoring verbal utterances.

Robot features. Robot features, i.e., how the robot should look, interact, behave, and how its environment should be shaped — were discussed at length. Participants discussed the robot's *form*, i.e. its appearance and embodiment. Participants were asked to consider 5 robots (Jibo, Pepper, Miro, Cozmo and Pleo) and rank them as most appropriate to least appropriate as a RC. Jibo received the highest points, followed closely by Pepper and Miro. Participants had differing opinions on whether they would prefer a humanoid or an abstract robot, but the consensus was that the robot's function should match its form. For example, participants noted that they would not expect the robotic dog Miro to talk, and would expect quite sophisticated, human-level *interaction* from a humanoid robot such as Pepper. Participants expressed that a two-way communication with a humanoid robot should be designed carefully, so that the robot is seen to understand and respond appropriately, and to avoid incongruous responses.

Robot *behaviour* was discussed from the perspective of being too *forceful* or *patronizing* vs. *not direct enough*. Participants wanted the robot to be encouraging, but they noted a balance is needed for this not to be patronizing. Participants explained that they want feedback on their practice as well as reminders, but these should be given on the appropriate level, and not be too intrusive. *Responsiveness* was also seen as important, as discussed in length as part of the robot's capabilities.

Robot advantages and disadvantages. Coaches and participants shared what they perceived to be the potential advantages and disadvantages of a RC, especially in comparison with a human coach, or a hypothetical mobile application with a well-being coaching function. While many participants noted that a robot could not replace a human as a primary method of delivery for well-being practices, it was seen to have the advantage of *accessibility* in comparison to a human coach (both in terms of timing and affordability). A robot's *physical presence* was seen to be a major advantage over a well-being practice being administered via mobile app or other digital means. However, *technological problems* such as poor internet connection or lack of software updates were seen to impede the potential advantage of accessibility.

Participants noted that the *lack of humanness* of a RC could be a major disadvantage. For instance, *unwanted robot behaviour*, such as incongruent responses when misunderstanding the user, could lead to participants feeling less connected or the session getting sidetracked. However, a RC could provide *reliability, consistency and uniformity* that a human coach could not. A coach noted that a robot would not get tired and could thus accommodate more participants, and would be more *consistent* and maintain *uniform* interactions due to not being affected by a daily routine, personal life, or a coach's personal factors such as nervousness or maintaining sufficient emotional distance.

Analyzing feedback data was seen as a potential advantage by the mindfulness coach. Participants also noted that a robot could record the practice, and adapt the practice according to data gathered. However, a disadvantage related to this was *invasive data collection*, where the SFP coach specifically mentioned that too much data such as biological signals could affect outcomes by making participants even more anxious. Participants were also concerned about *privacy issues*, although the SFP and life coach noted that a robot may actually provide more privacy if data is appropriately protected, since there is no room for human error.

Neutrality and anonymity was the final advantage of a potential RC perceived by the coaches. The SFP coach noted that a robot could be neutral in gender, and less intimidating than a human coach if it were smaller in size. The robot could also be *less judgemental*, as noted by the SFP coach

Table 2. Quotes from coaches (Ci, ($i = 1, 2, 3$)) regarding suitable well-being practices for a robotic well-being coach.

Well-being Practices	***Quotes from coaches***
Solution-Focused Practice	**The Miracle Question** C1: "The miracle question. You go to sleep and a miracle happens in the night. [...] The reason for the problem disappears and it solved the problem [...] but you were asleep and you didn't know the miracle that happened. What would be the first signs that something had changed, [and you] amplify what [the participant] said. [...] They'll say, 'This thing would happen' and [you ask] 'What else, what else? [...] So there are some questions that are very standard." C1: "[...] the words are very carefully chosen for a certain effect, so that could be very easily automated. It's already scripted basically [...]."
Meditation / Mindfulness	**Mindfulness of Breathing** C2: "[...] it's four stages. [...] You start with counting your breaths, after you breathe out, then breathe in. Then you stop, drop, [...] counting. You concentrate on the touch of the breath. Feel it. It's easy to instruct." C1: "[...] Breathing exercises (not just mindfulness or safe place visualisation) would be a good skill for a robot to be able to teach. It would have good pacing and could talk through the breathing script whilst also seeming to practise with you (without running out of breath which is hard as a human!)." **Loving Kindness (Metta Bhavana)** C2: "It's a bit more complicated [than Mindfulness of Breathing] as you need to [...] give [the participants] a bit more complex [information]. I'm sure you can do that because if a recording can do that, a robot might do that as well."
Life Coaching	**Mindfulness Exercises** C3: "These exercises [are already] recorded, audio with some relaxation exercises or imagination exercises, [and] guiding people through those." **Other Exercises** C3: "[...] Many of these exercises, whether they are [...] drawing, writing, answering questions. They're ready exercises so they can be recorded or they can be documented. And then the robot could easily show these in the written format or talk [to the participants]." C3: "[...] The Pepper robot had the screen there, so if it [were a] touch pad type of screen, a person could draw there, write there, [...] at least some words. There could be, in the same area where they meet the robot, [...] paper where they could write."
Cognitive Behavioural Coaching	**Fear Hierarchy** C1: "CBT is quite like you're teaching certain skills, like graded exposure might be part of CBT [...]." C1: "[...] Developing a fear hierarchy and helping someone. And track their progress. But I think you'd want it to be in tandem with a person." **Behavioural Activation** C1: "This when in particular people are depressed. You can do behavioral activation, which is a part of CBT, which is scheduling in activities that are meaningful for someone and tracking how that makes them feel. I could see that kind of activity scheduling working in a robotics context." C1: "But I think you would want it alongside a person who helped you generate ideas."

and meditation coach. The life coach also noted that a robot could be more neutral due to no bias through its *lack of life history*.

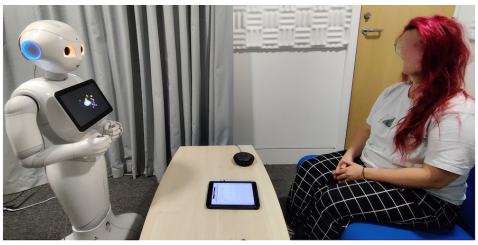

Fig. 4. Participant and Pepper interacting with each other.

4.2 Study 2

Our findings in Study 1 showed that the adaptability and emotional responses are important aspects for a robotic coach to deliver mental well-being practices. With the help of one of the mental well-being coaches who participated in Study 1 and motivated by those results, we designed the behaviour and interactions of a Robotic Coach (RC) to perform a brief, one-off positive psychology session with healthy adults as a proof-of-concept. We focused on the positive psychology practice given the expertise of the human coach who collaborated with us for designing this study. We aimed to examine the *robot capability* (see Section 4.1.5) of *responsiveness and adaptation*, specifically displaying *empathy* via the robot's *emotional expressions*, in response to participants sharing their experiences. This study is originally presented in [25], where we applied Continual Personalization to accomplish the adaptation capability.

4.2.1 Participants. In a one-session human-robot interaction study, 20 participants (12 females, 5 males, 3 not disclosed, aged 26.70 ± 3.68 years old from 12 different nationalities) recruited from the University of Cambridge student population interacted with the RC Pepper conducting Positive Psychology exercises over a 30 ± 11 minute session. The study consisted of three interaction rounds, each containing three exercises: 1) Two Impactful Things, 2) Two Things You Are Grateful For, and 3) Two Accomplishments. Each interaction round focused on the past, present, and future, respectively. The study examined the use of affect-based adaptation using Continual Personalization. The implementation of this adaptation is out of scope for this paper, and is described in more detail in a paper detailing the technical implementation of the robot and the corresponding interaction flow [25].

4.2.2 Setting. The study was conducted in a dedicated room at the University of Cambridge. It was equipped with two cameras facing the robot and the participant, in order to capture the interaction. The Pepper robot's on board RGB camera was used to capture the participants' facial affect. An external microphone was used to capture the participants speech, and Pepper used its on board speakers to communicate with participants. The participants were sat in front of the Pepper robot, with a table separating them in order to create a natural distance, and with the tablet containing the questionnaires to be completed during the study placed on the table (see Figure 4).

4.2.3 Protocol. The structure of the one-time interaction sessions can be seen in Table 3. Each interaction session consisted of a pre-interaction survey, introduction phase (with explanation of PP), three rounds, focusing on the participants' life in the *past*, *present*, and *future*, and a final post-interaction interview. The HRI script design was based on pre-existing positive psychology exercise literature, and was examined with a professional coach before the study, in order to mitigate any potential negative perceptions in the participants.

Items	**Duration**
Pre-interaction survey (writing)	5 min
Robot introduction	3 min
Explanation of PP	3 min
"Past" round	10 min
"Present" round	10 min
"Future" round	10 min
Conclusion	3 min
Post-interaction interview	8 min

Table 3. Structure of the study 2. Average times for each session were: past ($11 \pm 4 minutes$), present ($9 \pm 3 minutes$), future ($10 \pm 3 minutes$), and overall interactions ($30 \pm 11 minutes$). Originally presented in [25].

Before starting the interaction (pre-interaction survey), participants were asked to fill out few questionnaires that are described in details in Section 4.2.4.

Each interaction round (*past*, *present*, *future*) consisted of three exercises. Each participant conducted all three exercises three times, once for each interaction round.

i. *Two Impactful Things:* In this warm-up exercise, the robotic coach asked participants to talk about two impactful things or events that either happened in the past two weeks (*past*), happened or are expected to happen today (*present*), or are expected to happen in the coming two weeks (*future*). The robotic coach also asked the participants to think about why these events happened and how they made them feel.
ii. *Two Things the Participant is Grateful For:* This exercise focused on developing *gratitude*; a concept emphasised during Positive Psychology. Focusing on gratitude can increase positive affect, subjective happiness and life satisfaction [29, 37]. The robotic coach asked the participant to recall or imagine (depending on *past, present* or *future* interactions) two things that they felt or might feel grateful for.
iii. *Two Accomplishments:* In this exercise, the robotic coach asked the participant to think about *past, present* or *future accomplishments*, focusing on self-esteem, which has been applied to increase well-being and ameliorate depressive symptoms [42]. The robotic coach asked the participant to describe past, present or potential accomplishments, strengths the participant applied or may apply to accomplish these [69], and how these accomplishments make the participant feel.

After finishing all three interactions with the robotic coach (post-interaction interview), a semi-structured interview was conducted with the participants, documenting their perceptions of the robot and the positive psychology exercises. The interview structure can be seen in Appendix D.

4.2.4 Surveys. Prior to the interaction, participants filled surveys on a tablet (the IPIP-20 personality inventory, and the PANAS mood inventory). During the interaction, the participants filled the PANAS mood inventory after each interaction round (the past, present, and future), as well as the Godspeed and RoSAS questionnaires regarding their impressions of the robot. Participants also rated the statements "The robot understood what I said", "The robot understood how I felt", and "The robot adapted to what I said and did" after each interaction round on a 5-point Likert scale.

4.2.5 Data analysis. The data collated and analysed in the next section were collected through the semi-structured interviews with the 20 participants conducted immediately after each interaction

session. These interviews were analyzed with TA, the results of which are introduced in the next section.

4.2.6 Summary of Main Findings. Participants were asked about their opinions on the performance of Pepper [2] as a Positive Psychology well-being coach, what they thought worked well, and what they would change. The TA resulted in 10 major themes, with focus on robot features, capabilities, as well as the content of the Positive Psychology practices (as seen in Figure 5). A summary of the TA quotes can be seen in Table 4 and the complete list of results can be found in [25].

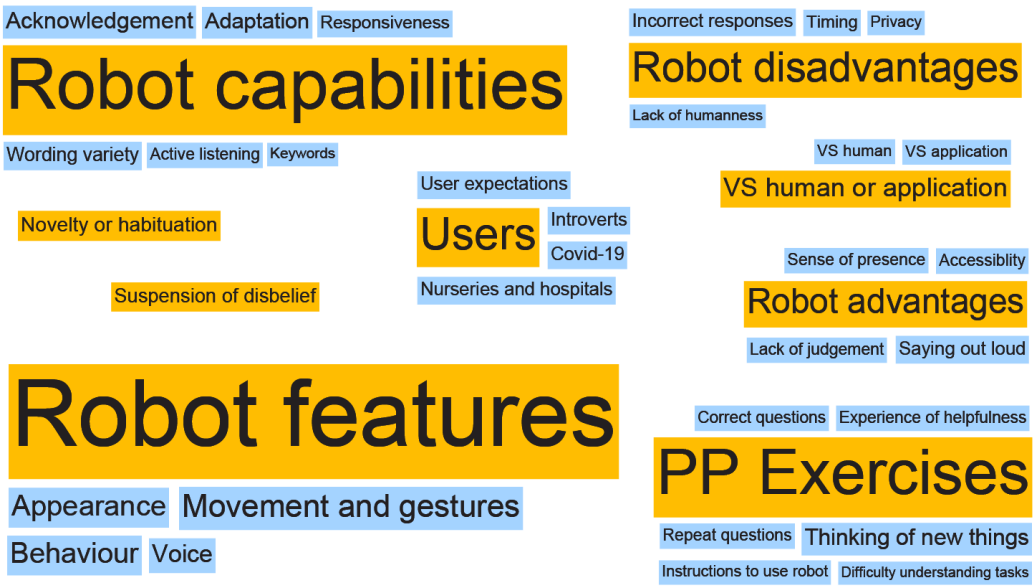

Fig. 5. Study 2: Themes defined in the TA are presented in orange, while codes related to these themes are presented in blue (best viewed in colour).

Robot capabilities. While participants enjoyed the content of the PP exercises, they would have preferred the robot to be more *responsive*. Participants noted that when they were disclosing their feelings (during task 1, Two Impactful Things), the robot could have "engaged more positively with their positive feelings". *Acknowledgement* of what the participants were saying, as well as a feeling of *actively listening* could be done via movement, as well as adding *wording variety* and having the robot respond differently to what the participants were saying (as opposed to repeating the same response every time). *Adaptation* was requested from the participants, specifically the robot having the ability to ask follow-up questions and give suggestions based on the participants' responses. Participants said the robot could pick up *keywords* from their speech, and adapt its utterances accordingly.

Robot features. Participants had mixed impressions of the robot's *behaviour*. Some perceived it as *supportive*, *compassionate*, *friendly* and as *giving space to talk* as well as *positive reassurance*, and that they felt *comfortable* and *heard* during the interactions. However, others noted that it was *robotic*, and even *condescending* when telling the participants "well done". One participant noted

[2]https://www.softbankrobotics.com/emea/en/pepper

Table 4. Selected quotes from participants about study themes gathered during the semi-structured interview lasting for ≈ 10 minutes. The quotes have been slightly edited for clarity.

Quotes about	*Quotes from participants*
Novelty and Habituation	**P17**: "I felt the last round [...] I got more used to the pattern [...], but it might have also been that it understood things more clearly." **P11**: "[...] it was weird to be talking to a robot. It's also the surprise when it first talked and moved, and then the first time it would say certain things that were kind of funny as well. It would say 'oh thank you so much for sharing that with me'. [...] I think once I got used to it, and it didn't take very long to get used to it, the weirdness wore off early."
Responsiveness and Adaptation	**P5**: "There were a few times when it did have a specific new response or it said 'That sounds like a great accomplishment' or 'You should be proud' and that was nice, I was like 'Oh thank you robot for listening to me'." **P13**: "[It had] good questions, but I don't think it adapted to anything I said [...] [It was] asking me questions that made me think productively. [...] [But] it felt more like the robot has a good script than it was engaging in active listening."
Exercises	**P7**: "At the end I think I felt more grateful. [...] I'm feeling more positive than I was before the start, in that respect it was effective." **P2**: "It was good that it explained the thought process behind the exercises, explaining why thinking of positive aspects was important. [...] I guess [it's] an exercise we could do ourselves, but it felt more interactive because the robot was asking you questions."

being stressed out when talking to the robot, since they anticipated a quick response from the robot, and wanted to think of a response quickly so that the robot would not move on prematurely.

The robot's *voice* was generally perceived well, with the *slow pace* being called patient and listening. A few participants found the *pronunciation* strange, with one noting being creeped out by it. However, another participant noted being positively surprised in that the voice wasn't as mechanical as they had thought.

The robot's *appearance* was generally regarded positively. It was called *cute*, *friendly*, and even *childlike and innocent*. The lack of facial expressions was perceived as a positive, with participants describing it as *approachable* and "taking in what I'm saying". The human form factor was perceived as "nice", however being at "the high end of humanoid before being too creepy".

Robot advantages and disadvantages. Participants made favourable comments on the robot in comparison to a hypothetical mobile app designed for the same set of exercises, noting that *speaking out loud* was in itself helpful, as well as the *sense of presence* that can be provided by its physicality. Advantages in comparison to a hypothetical human coach performing the same set of exercises were that the robot had a *lack of judgement*, where it might be easier to talk to a robot due to it not being as reactive as a human, as well as *accessibility*, where a human coach is limited to working hours and a robot can serve more people in a day.

However, a robotic coach was also perceived to have disadvantages. Participants noted that the robot had problems with *timing*, where it sometimes interrupted participants before they had finished talking, and paused for too long between responses. These were perceived to be a distraction, and left participants feeling like they were being hurried up. The robot would also give *incorrect responses* on occasion, robot's emotion recognition system made incorrect estimations of the emotion expressed by the participant. This made participants feel misunderstood. Another

disadvantage of the robot was perceived to be *privacy issues*, where participants were sensitive being recorded, and chose what to talk about on that basis. Finally, participants noted that while it could be an advantage in some cases, the *lack of humanness* of the robot could be limiting in the depth of information it can ask, and couldn't give organic responses.

Positive Psychology exercises. Participants especially made several comments on the Positive Psychology exercises, noting that they were *helpful* and made them reflect, particularly about a positive future. They noted feeling more grateful and positive after the exercises, and that the robot was useful for guiding them through the exercises. Particularly, the robot asked the *correct questions*, and that it explained the thought process behind the exercises. Participans noted that they *thought of new things*, e.g. why they felt grateful, and what they felt grateful for. Participants noted that the robot could be improved by having the ability to *repeat questions*. Some participants didn't hear or understand the robot the first time it explained a task, and wanted to be able to go back and ask for clarification. Some participants also had *difficulty understanding the tasks*, noting that they weren't sure of the exercise structure. Here, too, it would have been helpful to be able to ask the robot for clarification. Some participants also noted that they wanted more *instructions on how to use the robot*, e.g. how loud they should speak and how long they should wait before responding.

Users. Participants talked about different aspects related to the *users* of the robot. Specifically, they noted that the *expectations* of the users coming to interact with the robot would shape the interactions. One participant noted that people who had not done therapy before, and thus did not have any expectations, could benefit from the robot the most. Another noted that it works on the basic level, but if a person had a bad day and was expecting more support, the robot could be a disappointment. One participant noted that the robot might be especially useful in the aftermath of *COVID-19*, and that it could be useful for people isolated in homes due to the pandemic, or even people who are just socially isolated but would want a channel to open themselves back up to society. Other participants concurred with this, noting that the robotic coach could be particularly helpful for *introverts* that felt shy or awkward with humans, and could feel more confident with a robot. Finally, one participant noted that the robot could be useful in *elderly care homes and hospitals*, where the small interactions with a robot could make a huge difference where nurses weren't available, especially for the older generations who are alone.

Phenomena. Two Human-Robot Interaction related *phenomena* were described by the participants. They are two separate phenomena, but are grouped here under the label *Phenomena* since they were separate from other themes selected in the TA. *Novelty and habituation* effects seemed to have an effect on participants' perceptions. In relation to the exercises themselves, one participant described understanding the exercises more clearly in the last round, as they had gotten used to them. However, another noted that they came to predict what would happen after the first round, and weren't reflecting as much anymore after that. Participants also described evolving perceptions of the robot, e.g. being initially surprised by it and laughing. Another described being initially creeped out by the robot's smile and some arm movements, but easy to get used to. *Suspension of disbelief* was the other HRI related phenomenon identified in the TA. Suspension of disbelief is a concept discussed in HRI [35] as the person's willingness to suspend their disbelief of what is living in physical social robotics, and how this willingess may affect people's perceptions of such social robots. The concept originally relates to an audience's willing participation in a fiction [82]. Participants did not explicitly mention suspension of disbelief by name, however, many participants talked about feelings of strangeness and being conscious of the experience.

4.3 Study 3

As this study is not presented in a prior paper, we present it here in more detail than Study 1 (Section 4.1) and Study 2 (Section 4.2).

The last step of the iterative design process consisted of another user-centred design study (Study 3) involving participants who interacted with the robot in Study 2. Study 3 was a group discussion aimed at examining what would motivate participants to use a robotic well-being coach in the long term, as well as what they would like to see from a robotic coach of Solution-Focused Practice (SFP).

4.3.1 Participants. We recruited 2 well-being coaches (one of them had already taken part to Study 1 described in Section 4.1), specialized in Solution-Focused Practice and Mindfulness respectively, as well as 3 prospective users who had tangible experience with the robot due to having attended the Positive Psychology exercises conducted by Pepper (as presented in Section 4.2 and in [25]). Participants were 3 females and 2 males aged 32.8 ± 9.1 from 4 nationalities (British, Brazilian, Portuguese and Romanian), and 2 of them have some experience with the humanoid robots, while the other 3 had no to little experience with them.

Fig. 6. Two participants joined the focus group via MS Teams, and the other 3 (one of them is not depicted in the picture because she was out of the camera frame) joined in-person.

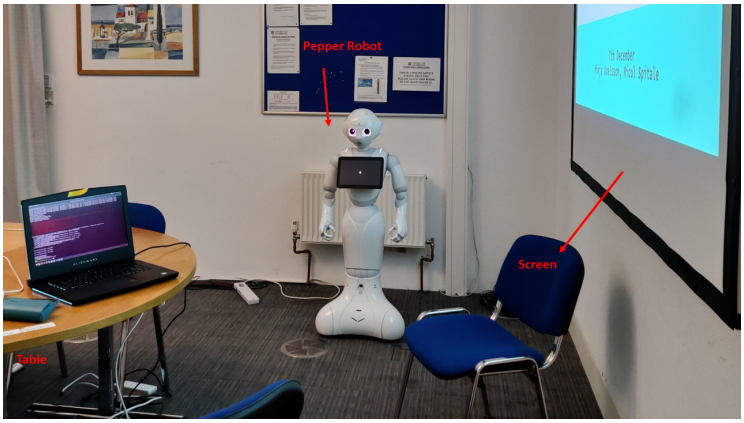

Fig. 7. Setting of the focus group: a table around which participants were seated, a Pepper robot that performed a Positive Psychology interaction, and a screen that displayed the information required during the discussion.

4.3.2 Setting. We conducted this study in a conference room of the University of Cambridge following an hybrid format. Three of the participants were present in-person while the other two

joined the study remotely via MS Teams as shown in Figure 6. The in-person participants were seated around a table. The room was equipped with a projector that was used to display slides that contained information for the discussions, accessible to the remote participants in an online format. We placed the robot Pepper by the screen for a demonstration of the Positive Psychology exercises. Figure 7 depicts the setting of the study.

Items	**Duration**
Demo of Positive Psychology with Pepper	5 min
Video of Mindfulness conducted with Pepper	5 min
Presentation of SFP by coach	5 min
Discussion after demo	25 min
Editing SFP script in groups	45 min
Ending discussion	20 min
Post-discussion questionnaires	5 min

Table 5. Structure of the focus group discussion

4.3.3 Protocol. The structure of the group discussion is detailed in Table 5. First, we introduced ourselves and explained to the group the main goal of the session. Then, we asked one volunteer among the in-person participants to interact with a Pepper robot to demonstrate how the robot can deliver a Positive Psychology well-being session. All the participants watched the demo that lasted for about 5 minutes. The robot asked the participant what they were grateful for, and what they had recently accomplished. A short discussion was held about the demo just experienced, prompted by questions we prepared (e.g., "Does the way the robot looks and talks feel right to conduct these exercises?") displayed on the screen for about 25 minutes.

After this, participants were also shown a video of Pepper conducting a Mindfulness session [17]. The coach present in-person introduced the concept of Solution-Focused Practice, as well as the main objectives of the practice, in order for all the participants to have initial knowledge on the topic. The coach answered the participants' questions related to the topic. The three demonstrations (Positive Psychology, Mindfulness, and SFP) were intended to give participants information about the range of well-being practices available, and to spark their imagination on what it would mean for a robotic coach to conduct them.

We then split participants into two groups (one of 3, and the other of 2 people), and we assigned the well-being coaches respectively to each group. The members of each group were asked to collaborate with each other to discuss the use of robots in well-being practices, as shown in Figure 8. Specifically, we provided them with a collaborative text document that contained a SFP dialogue script that had been transcribed based on sessions previously conducted by the SFP coach. Participants then edited this script based on how they would like a robotic coach to deliver SFP sessions (see the SFP script given to participants for editing in Appendix E). The group members were asked to edit the script and add comments to improve the sessions to be delivered by the hypothetical robot. We also supplied them with the online tool *Miro* incorporating additional questions as well as portions of the Social Robot Co-Design Canvases [6, 9].

After 45 minutes, we began a group discussion on the topic of a robotic coach conducting SFP sessions, based on the edits the participants had made to the robot script. Again, we prompted them with a list of questions (e.g. "What would motivate you/your clients to start using a long-term

Fig. 8. Participants are discussing the SFP script during their group activity.

robotic coach?") projected on the screen to facilitate the 20-minute discussion. Finally, we asked the participants to fill questionnaires to evaluate robot capabilities (described in detail in Section 4.3.4).

4.3.4 Surveys. We asked participants to fill out two surveys: a questionnaire with open-ended questions (e.g., "Refer to what you've put in the Miro board: What robot behaviours did you agree on with your pair?" or "Refer to what you've put in the Miro board: What 3 things would make you/your client most de-motivated to use a robotic mental well-being coach?") to evaluate the motivation to start and keep using a robot well-being coach, and a demographic questionnaire.

4.3.5 Data analysis. We analyzed the data from the group discussions using TA following the method described in Study 1 (Section 4.2).

4.3.6 Main Findings. Figure 9 depicts the themes identified in the thematic analysis. The following subsections report the themes and related findings, with quotes from the transcripts summarised where necessary (reported in Table 6, 7, 11, 12, 8, 9 and 10).

Table 6. Quotes from participants (Pi, ($i = 1, 2, 3$)) and coaches (Ci, ($i = 1, 2$)) regarding *robot capabilities* theme.

Robot capabilities	***Quotes from participants and coaches***
Turn-taking	P1: "[...] you don't know if the pause is like, it's waiting for you to stop speaking so it can carry on, or if it's just you know, broken or something." P3: "We don't understand whether our answers are being processed or not, and so there has to be some sort of cue that the robot is thinking or processing our information, and did not just malfunction and die."
Empathy	C2: "[...] I also feel at the moment Pepper is lacking compassion. Being there with the person and sharing in their suffering and empathising." C1: "[Giving empathy at the person's level], it's better [that the robot goes] just slightly above. Like if they're giving a little bit of hope, you just give a little bit more. You don't go wildly above, otherwise it feels really jarring. I mean when someone speaks quietly, we speak quietly. I think the robot could do that, would just [speak] slightly louder and I think that's what people do. Like you don't want your coach to sound depressed, but you also don't want them [to be jarring], so 'I'm really jolly when you're down here', yeah?"

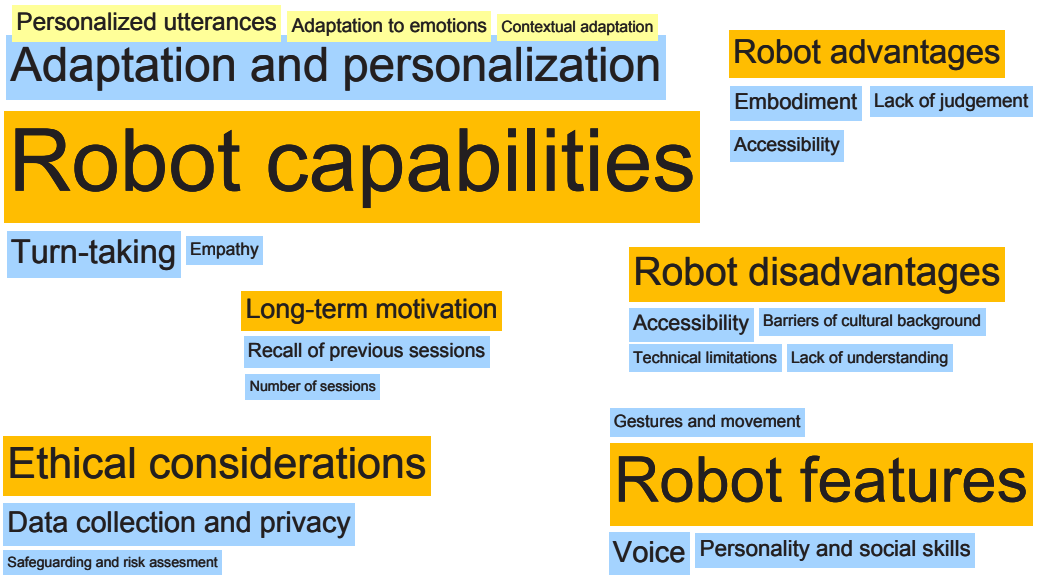

Fig. 9. Study 3: Themes defined in the TA are presented in orange, while codes related to these themes are presented in blue, and sub-codes are presented in yellow (best viewed in colour).

Robot capabilities. Robot capabilities are seen in Table 6, as well as a separate table for the capabilities sub-theme *Adaptation and Personalization* (Table 7). Coaches and participants thought that the *turn-taking* (i.e. improving signals of when the robot is listening and when it is about to speak) of the dialogue is crucial for a successful human-robot interaction. They reported that the feeling of being listened to and understood needs to be enhanced. They also believed that the gestures could be misleading and not clear enough indicators of turn-taking, suggesting that cues for this need to be designed further. Another important robot capability was conveying *empathy*, as currently the robot was *lacking compassion* as perceived by participants. The SFP coach suggested empathizing *at the person's level*, in order to avoid *jarring* emotional expression. For example, if the person were feeling very low and speaking quietly, the robot should speak just slightly louder and give slightly more hope, rather than being highly expressive and joyful.

During the focus group, participants agreed on the fact that the most important features of a robot coach for well-being should be *adaptation and personalization* to the user (see Table 7). We grouped adaptation and personalization as a single theme because participants used those terms as synonyms during the discussion. Coaches and participants believed that the robot should have *personalized utterances* to each user to avoid the feeling of a "mechanical" interaction, *adaptation to the user's emotions* for the coaching to be more successful, as well as *contextual adaptation* to adapt to the context that the robot is being used in. Some of the participants pointed out that the robot should personalize the wording of its utterances to each user, even across sessions. However, the SFP coach highlighted that its personalization could be counter-productive in some cases, as SFP sessions have a set, scientifically evaluated and approved structure, that should be followed in order to attain the well-being benefits from the sessions. This indicates that the suggestions made by prospective users (who are laypeople uneducated in psychology) and professional coaches (educated in the underlying mechanisms of the well-being practices) may be in conflict. Careful evaluation is needed in determining when following each suggestion is appropriate, i.e. when

Table 7. Quotes from participants (Pi, ($i = 1, 2, 3$)) and coaches (Ci, ($i = 1, 2$)) regarding *robot capabilities: adaptation and personalization* theme.

Adaptation and personalization	***Quotes from participants and coaches***
Personalized utterances	C1: "Maybe add some personalised comments in the middle about what the user said specifically?" P3: "Making it personalized, [...] picking on what people are specifically saying. [...] I don't want to know that my interaction with the robot was exactly the same, or [had] the same answers, as someone who had a completely different issue. So how can you make it personal?"
Adaptation to emotions	C2: "Just that the language should be adaptive and based on the responses of the participant, e.g. the response to a [mood score of] 2 should be very different from one to a 10." P2: " I was thinking that we humans very rarely give a straight answer of 'I'm feeling happy'. Normally we'd balanced both [ends of the emotional spectrum] in the same sentence like 'oh, this thing happened'."
Contextual adaptation	C2: "It would be great [for it to adapt to context] but I think this would be really tricky. I would prefer that they respond and adapt to what the person is saying and their current mood/needs." P3: 'Things like language spoken, is it a special event, day, season, religious context, location where the robot is based..."

to follow the coaches' suggestions to preserve coaching efficacy, and when to insert adaptation as suggested by the propsective users in order to create an experience of *adaptation* and *active listening* in the robotic coach. The desire for adaptation and personalization also conflicted with the participants' and coaches' concern about privacy and data collection, and they were apprehensive about storing data for long-term interactions. This is discussed further in Section 4.3.6.

Table 8. Quotes from participants (Pi, ($i = 1, 2, 3$)) and coaches (Ci, ($i = 1, 2$)) regarding *robot features* theme.

Robot features	***Quotes from participants and coaches***
Voice	P1: "I'll be quite surprised if Pepper's prosody will be able to make this long hypothetical bit [[in reference to Miracle Question in SFP]] sound natural." C1: "Good point of it being hard to follow a long block of monotonous text from a robot, and [that you] can't ask it to repeat."
Gestures and movement	C2: "[..] the hand gestures seemed fine, although I agree that a tilt of the head would be really natural as well." P3: "[it was] uncanny, I was like this is a robot, but the hands look too human. I don't know what to do with your movement. I did not understand."
Personality and social skills	C2: "I think pepper should use more encouraging, motivating language that reflects the person's strengths — resilience, tenacity, desire to move things forward and change things about their lives, and willingness to engage in an exercise like this that is in their benefit." P3: "[...] making the robot have certain characteristics means that the participant will feel empathy towards it, doesn't it? What if the robot describes being in an emotional state, that the participant feels conflicting with their own state? [...] Perhaps friendly is a good personality?"

Robot features. Participants discussed what features a robotic well-being coach should be endowed with (see Table 8). The Mindfulness coach emphasized that the robotic coach should display *empathic behavior* and compassion, and underline the user's *strengths*. Coaches and participants agreed that it would be important for users to feel that the robot was able to *understand* them. Some participants noted that the synthesized voice of the robot lacked *intonation and prosody*. They believed that the *flat* and *monotonous* voice could lead to an unnatural feeling during the interaction. They suggested that the robot should not say very long sentences in order to sound more natural.

Some participants experienced an uncanny perception during *robot's gestures*, calling for more *acknowledgement* gestures, i.e. nodding during the participant's speech. Participants believed that the robot's *personality* should be neutral, but still warm and friendly. The robot should show positive behavior without being over enthusiastic during the coaching.

Table 9. Quotes from participants (Pi, ($i = 1, 2, 3$)) and coaches (Ci, ($i = 1, 2$)) regarding *robot's advantages* theme.

Robot's advantages	***Quotes from participants and coaches***
Embodiment	P3: "What the robot does bring is an embodied [..] it's the feeling that you're interacting with with a material form [..] can bring it does bring a different reaction than if we were just speaking to a, uh, a TV or or something flat." C1: "If it even vaguely looks human, they would be more likely to share [their] thinking, it's not like a device."
Lack of judgement	P1: "Some clients might feel more at ease with technology than with other humans." P2: "I would in some cases fields more opens it, just tell pepper things that other humans 'cause they would have less judgement." C1: "Some of them might have preferred talking to a robot, knowing that it's not going to go anywhere like this. Yeah, there's no trust issue." P3: "[...] caters to people uncomfortable with sharing their personal situation with other people."
Accessibility	C2: "Not a person, but you know someone that you can talk to [...], has a friendly feel to it. I can imagine if I had one in my room I might well [...] get into the habit of chatting with it, but I don't know whether the benefits would be great enough versus a phone app, or a zoom call." P1: "The robotic coach will also have more availability than a human."

Robot advantages and disadvantages. Coaches and participants thought the robot might have certain advantages (Table 9) and disadvantages (Table 10), particularly discussing these in comparison to a *mobile app* or a *human coach*. Some participants saw the robot's *embodiment* as an advantage in comparison to a mobile app, while the Mindfulness coach and some participants saw that the *barriers of the technology* were limiting what the robot could accomplish, and thus would not see much difference in using the robot in comparison to a mobile app. The robot's *lack of understanding* was also seen as a limitation, in particular the SFP coach remarked that people would have to sometimes repeat themselves loudly when talking to the robot. However, one participant remarked that "It doesn't matter what we say to the robot, but perhaps it matters what I say to myself", indicating that the robot could be useful even with limited understanding.

The robot's *lack of judgement* was seen as an advantage, as it could help users *trust* the robot more. Participants noted that people who were uncomfortable with sharing personal situations with other people might find the robot useful. On the other hand, barriers related to *cultural background* could reduce this advantage. P3 noted that in some cultures, emotions are not openly expressed,

Table 10. Quotes from participants (Pi, ($i = 1, 2, 3$)) and coaches (Ci, ($i = 1, 2$)) regarding *robot's disadvantages* theme.

Robot's disadvantages	***Quotes from participants and coaches***
Technical limitations	C2: "[..] prompting people to reflect on what they're grateful for, is really beneficial. I'm not sure for me what Pepper is adding to that process, versus an app on your phone that might just pop up at a certain time and say, tell me what you're grateful for. [..] What does Pepper bring that is beyond just the auditory stimulus to reflect?" P1: "I think that there is still quite a lot of programming and technological [work to be done] [...]. It's increasing the type of reactions, understanding the words people are saying, not requiring people to repeat themselves, giving some sort of continuing body movement. I agreed, the hands were very creepy when I did the exercise." P3: "Maybe it doesn't make any difference that it's a robot, or we could do the same by any other virtual means."
Lack of understanding	C1: "You're having to repeat, and repeat loudly, when you're doing an exercise [in order for the robot to hear you]." P2: "It doesn't matter what we say to the robot, but perhaps it matters what I say to myself."
Barriers of cultural background	P3: "[...] it touches upon all of these very tricky feelings and emotions. That in a lot of cultures, they're not to be openly expressed. Even to a robot, [...] you don't share it. And so the robot works in specific contexts, it does. But then in other sociocultural contexts, the robot will not be able to. People won't give the robot the information that they're supposed to give. It's not culturally acceptable to do that." C2: "[...] we just need to make make it clear that it may not be adaptable to other cultures and people."
Accessibility	P2: "[..] money issues." C2: "The barrier to using it would be the cost and availability and access to Pepper, which would mean that the benefits would have to be really quite big [...]"

and this could apply also to expressing them to a robot. As noted, the deployment of such robots should be carefully considered in light of the cultural context of its environment.

Accessibility of the robot was seen as both an advantage and disadvantage. Some of the participants believed that cost and availability are contributors to the unpopularity of robots currently, and that they were unsure whether a robot's benefits would outweigh such barriers. On the other hand, some participants noted that having the robot in their room would encourage them to use it, and that a robotic coach could have more availability than a human coach.

Motivation over time. The group was also prompted to discuss what would make them motivated to use a robotic well-being coach over longer periods of time (see Table 12). Interestingly, the SFP coach clarified that the duration of SFP coaching actually depends on the coachees, and how many sessions they needed and benefited from. The coach mentioned that the coaching has succeeded when it is no longer needed. For example, some people could need only a single session to feel better, and others could need 10 sessions before considering the well-being coaching successful. However, other well-being practices such as Mindfulness practices are different in that they can be practiced every day, without the intervention targeting a particular problem, and can continue indefinitely. Participants discussed that in the case users wanted multiple sessions, i.e., a longitudinal interaction with the robotic well-being coach, an important robot capability is the robot's memory of the

past sessions. The robot should recall how the users felt during past sessions, because it could be beneficial and rewarding for the user to check their progress. This would also help facilitate the perception that the robot was *acknowledging* and *actively listening* to the participants. Participants also stated that the robot should avoid *repetitive behavior* in long-term interaction in order to avoid the perception of the robot being mechanical in nature, which might create frustrations in the user.

Table 11. Quotes from participants (Pi, ($i = 1, 2, 3$)) and coaches (Ci, ($i = 1, 2$)) regarding *ethical considerations* theme.

Ethical considerations	***Quotes from participants and coaches***
Data collection and privacy	C2: "We need to be very transparent, very upfront, and the conversation about you know what, if anything is being recorded and transmitted." P1: "I didn't want to share something extremely intimate [in a group setting with the robot]. Yeah, but I was trying to say something that makes sense as an answer, but it's not at the same time exposing myself too much."
Safe-guarding and risk assessment	C2: "There may be at risk of suicidal ideation or self harm or and I just feel that [...] we need to risk assess first before we get into the issues that that person wants to talk about it." C2: "If Pepper comes across someone who is in crisis, either suicidal, self-harming, or in danger. What are the safeguarding processes? Is there a way Pepper can check if people are at risk before the conversation starts?"

Ethical considerations. Participants discussed several ethical considerations related to robotic well-being coaches, as presented in Table 11. The participants' main concern when interacting with a robotic well-being coach was *data collection and privacy*. During the discussion, the Mindfulness coach highlighted that *transparency* is a key factor for a successful coaching session: the robot should state upfront if it is collecting or recording any data, in compliance with GDPR (the General Data Protection Regulation of the European Union [98]). However, transparency could also affect the efficacy of coaching. The volunteer participants during the Positive Psychology session noted that they were considering what they should share with the robot, as they were conscious that other people could hear them. One of the discussion groups noted that privacy could especially be a concern in countries with totalitarian regimes, where expression of negative emotions or concerns is not socially accepted, and may even be surveilled. This indicates the importance for cultural consideration when deploying such robots, and the responsibility of roboticists to ensure that sensitive data is appropriately protected.

Another ethical consideration highlighted by the well-being coaches was the importance of *assessing risk and safeguarding* during well-being coaching. Coaches expressed that the robot would not be appropriate to be used with clinical patients experiencing severe distress (e.g. suicidal ideation). The discussion converged in the viewpoint that users of such a robot should be pre-screened by human practitioners for any warning signs, as a robot may not be able to assess these types of situations.

5 QUALITATIVE ANALYSIS: CONVERGENCE AND DIVERGENCE OF THE STUDIES

The three studies presented in this paper converge and diverge in different themes defined with the TA. As all studies were part of an iterative design process, each study built on the conclusions of the previous ones. Due to the flexible and open nature of the study procedure, differences and convergences emerge between the feedback gathered from the stakeholders. In the next sections,

Table 12. Quotes from participants (Pi, ($i = 1, 2, 3$)) and coaches (Ci, ($i = 1, 2$)) regarding *motivation over time* theme.

Motivation over time	***Quotes from participants and coaches***
Number of sessions	C1: "Interventions should be as brief as they need to be or can be, so I wouldn't want it to be used indefinitely. [...] It depends on the person. A person could need a single session or 100 sessions."
Recall of previous sessions	C2: "There's no reminders of how well I'm doing or how positive it is that I'm engaging in this thing — in this process that could be to my benefit [...]. [Thinking] 'I am doing quite a good job', 'I'll keep doing it', so [adding more] strength language in there." P1: "Remembering what happened previously and saying, 'oh, you are feeling better than you were last week'?"

we will examine the *convergence* (i.e. similarities) and *divergence* (i.e. differences between) study results, in order to highlight both complementary findings, as well as points of departure and the reasons for these departures [96]. All themes and their occurrence across the studies are presented in Appendix A.

Table 13. Comparison of 3 studies

	WBP	**Coach**	**Duration**	**Data collection**
Study 1	BSFP; Mindfulness	Human, first person	Long-term (four 20 min sessions over 4 weeks)	Individual interviews, group discussions
Study 2	Positive psychology	Robot, first person	Short-term (one 30 ± 11 min session)	Individual interviews
Study 3	Positive psychology; Mindfulness; BSFP	Robot, first/third person	**PP:** Short-term (one 10 min observed session), in some cases took part in Study 2 as well; **Mindfulness:** Video demonstration; **BSFP:** Introduced by coach	Group discussion

5.1 Convergence

All three qualitative studies detailed similar capabilities and features of a robotic well-being coach, as well as what advantages and disadvantages it may have, especially in comparison to a human coach or a mobile app delivering the same exercises. This section details which of the sub-themes within these themes were expressed similarly across studies.

5.1.1 Robot capabilities.

Acknowledgement and active listening — Study 2 and Study 3 emphasized the robotic capabilities of *acknowledgement* and *active listening*, and how these capabilities could be improved

in order to enhance the user's experience and *feeling of being listened to*. Participants noted that these capabilities could be expressed through *verbal utterances* by using *wording variety*, *keywords* from the participants' speech, *paraphrasing* and *summarising*, as well as *phatic expressions*. *Body language* could also be used to convey *backchanneling*, with the robot nodding, leaning in, or tilting its head.

Expressing empathy — All three studies mentioned the importance of the robot *expressing empathy*, with Study 3 specifically going into detail on what level of empathic expression is appropriate (see Section 5.2.3). Empathy was seen to be related to the *feeling of being listened to*, and could on a more sophisticated level include *emotional adaptation*.

Adaptation and personalization — *Adaptation and personalization* were considered important in all three studies, with Study 3 especially going into detail on *personalized utterances*, *adaptation to users' emotions*, and *contextual adaptation* (see Section 4.3). Participants felt that adaptation and personalization (terms which most of them used interchangeably) could help them *engage* in the practice. However, the appropriate level of this adaptation was not agreed upon by all participants (see Section 5.2.3 for discussion).

5.1.2 Robot features.

Form — Robot features were discussed in a broad and hypothetical sense (without reference to any specific robotic implementation) in Study 1, and more specifically in Study 2 and Study 3, where there was a specific robotic application to evaluate and improve. While the preferences for the robot's *form* — i.e. humanoid and abstract robots — diverged in Study 1 (see Section 5.2.1 for discussion), the humanoid robot Pepper received mainly positive feedback during Study 2 (with some participants still noting they would prefer a more abstract robot).

Voice and movement — The robot's *voice* and *movements* were mainly discussed during post-interaction interviews in Study 2. Pepper's default voice, which was used during these studies, received positive feedback for its slow pace and friendly tone, and the movements were generally received positively, but needed improvement with regards to *expressing acknowledgement* and *less repetition*. Additionally, participants in Study 3 noted that Pepper's default voice could be slower and use more variation in prosody, in order to better engage the user and be more understandable.

Behaviour — The robot's *behaviour* was mainly examined via post-interaction interviews in Study 2, and via group discussion in Study 3. In both studies, participants were focused on improving the capabilities of *responsiveness* and *adaptation* in the robot's behaviour (see Section 5.1.1). Outside of these capabilities, participants in Study 1 discussed the robot's behaviour from the perspective of being *too forceful or patronizing* versus *not being direct enough*. In Study 2, the robot's behaviour was widely regarded as *friendly* and *supportive*, with only a few participants mentioning that the robot could at times come off as patronizing. Some participants in Study 2 found it positive that the robot did not express too much emotion. In study 3, participants also emphasized that the robot should be friendly and compassionate, as well as not too emotionally expressive.

5.1.3 Robot advantages.

Lack of judgement — The robot was also seen to have a *lack of judgement* in comparison to a human coach, across all three studies. In Study 1, two coaches noted that the robot would have this as an advantage over people, with the Life Coach noting that their life experiences could bias them to attempt to solve the coachee's problems in a non-productive way. Participants in Study 2 noted that the robot Pepper not having facial expressions was actually a strength, as it could take in what the participant was saying without reacting and being judgemental. In Study 2 and Study 3, participants brought up that users who were not comfortable sharing their feelings or

thoughts with other humans could make use of robotic coaches. Future work should explore how in particular introverted or socially isolated people could benefit from a robotic coach.

Presence and embodiment — The robot advantage of *embodiment*, i.e., *having a presence*, was seen in both Study 1 and Study 3. Participants noted that having something physically present, instead of using, e.g., a mobile app, could help participants feel *more engaged* in the well-being practice. In Study 1, participants noted that the presence of a robot would also act as a *visual reminder* for the practice. In Study 2, participants noted that one of the major advantages of the robot in comparison to a mobile app was *talking out loud*. Since the participants regarded the robot as a conversation partner, they noticed that even saying their thoughts out loud was helpful, even if the robot did not always understand what they said. This same sentiment was expressed by a participant in Study 3.

5.1.4 Robot disadvantages.

Lack of humanness — *Lack of humanness*, i.e. a robot's lack of human-like communication, was a disadvantage mentioned in Study 1 and 2. This is the other side of the picture with the robot's advantage of *lack of judgement* — while a robot can provide a judgement-free well-being platform, sometimes those reactions can be desired and even deemed necessary by participants. Specifically, participants wanted to be able to dynamically correct *incorrect responses* of the robot during conversation, i.e. resolve misunderstandings. Additionally, during Study 2, participants wanted to be able to *ask for clarification*, as well as *receive suggestions* based on their answers to the robot's questions. While these capabilities could be added to a robot, the level of intuitiveness that people have (e.g., the coach being able to recognize a look of confusion on a person's face, and providing clarification due to this) may not be achievable with a robot. This is related to Study 3, where participants noted that *technical limitations* could demotivate them from using a robotic mental well-being coach.

5.1.5 Ethical considerations.

Privacy and data collection — *Privacy* was the main perceived ethical concern a with robotic mental well-being coach in Study 1, 2 and 3. Especially in Study 2, a participant with robotics experience noted that they were conscious of what experiences they were sharing with the robot since they knew the data would be reviewed by human researchers later. The same participant repeated the same concern in Study 3. During Study 3, the Mindfulness Coach emphasized that the robot should not collect or store any data — however they also said that the robot should be able to adapt its behaviour to each participant according to their wishes. Due to this conflict, the disadvantage of privacy concerns is particularly tricky, as participants may not understand that full preservation of privacy and capabilities wished for from the robot may be incompatible. On the contrary, the SFP Coach involved in Study 1 noted that a robot may actually provide more privacy than a human, as in some situations a human coach may be required to disclose important information to other people, or may disclose something due to human error (as also noted by the Life Coach in Study 1). To mitigate this concern, informing participants of what data is stored and how it is used could be applied. This is discussed further in section 6.

Safeguarding and risk assessment — Safeguarding and risk assessment came up as an ethical consideration in Study 1 and Study 3. In Study 1, participants noted that the robot should have a protocol for if a user were in distress. In Study 3, the Mindfulness coach in particular noted that users should ideally be pre-assessed before interacting with the robot, and directed to other resources if they were in crisis. The coach and some participants remarked that the robotic coach

would only be appropriate for people who did not have significant mental health challenges, and were rather seeking to improve their well-being.

5.2 Divergence

Divergences in themes between the studies may be due to the different framing of all of the studies (i.e. different topics of discussion, different data collection methods, different well-being practices being discussed). Divergences between participants' opinions also emerged within the studies, possibly as a result of different perspectives (participant versus coach) or different backgrounds (SFP coach versus Mindfulness coach, or participant in technology field versus anthropology field). Additionally, the three studies had different approaches since they were conducted in consecutive phases of the iterative design process: Study 1 was a hypothetical, broader discussion of the potential of robotic mental well-being coaches; Study 2 examined a particular implementation of Pepper as a Positive Psychology coach; Study 3 gave an overview of three robotic well-being coaches and asked participants to directly edit a script for a SFP robotic coach, and to discuss their opinions (see Table 13). This section first reviews the divergences in opinion, and then the divergences in the emphasis of different themes due to the different framing of studies.

Divergences in opinion

5.2.1 Robot form. Robot form was discussed in detail in Study 1, while in Study 2 and 3 the robot's form was assumed to be Pepper. In Study 1, participants disagreed with each other about the robot's form, with some preferring a more abstract robot and others a more humanoid robot (see Section 4.1). However, all agreed that the robot's form should follow its function — i.e. that they would expect a more humanoid robot to have more human-resembling conversational capabilities.

While participants had differing opinions on preferred robot form in Study 1, Pepper's form was largely accepted during Study 2 and Study 3, where it was not explicitly a topic of discussion. A few participants called for a more abstract robot during Study 2, and some had specific criticisms of Pepper's form (the hands were creepy, or the light in Pepper's eye was a distraction). However, participants did note during Study 2 that they desired more human-like conversational qualities from the robot (i.e. responsiveness and adaptation) due to its human-like form. Similarly, while the form of Pepper was not explicitly discussed during Study 3, participants focused mainly on designing responsiveness and adaptation features for a robotic coach, after being presented the robot Pepper as an interaction platform.

5.2.2 Adaptation and privacy. During Study 3, the conflict between people wanting adaptation and privacy simultaneously was discussed in detail. This concern became apparent when participants were editing the SFP script of the robot, helping them understand that wherever they wanted the robot to adapt or personalize specifically to data they had supplied during the current session or the previous session, the robot would have to record such data. This concern was not discussed in Study 1, as the study was more about a hypothetical robot. In Study 2, participants also called for more adaptation, and while they were concerned about privacy, they did not state as strongly that the robot should not collect any data. Rather, they emphasized that the participant should be well informed about what data is collected (which had taken place before the study).

5.2.3 Emotional adaptation. In Study 1 and Study 3, the SFP coach held the view that it would be better for the robot to not adapt to a person's emotional experiences too much, due to the risk of being jarring if the robot got the reactions wrong. The coach held that the robot could be helpful even with minimal emotional adaptation, while retaining a *friendly and neutral personality*. However, during Study 3, the Mindfulness coach as well as some participants said that the robot in its current

state was not compassionate enough, and more *adaptation to emotions*, as well as *expressions of empathy*, were necessary for the robot to be useful. These participants had *expectations* toward a robotic coach that can be difficult to achieve due to *technical limitations*. Participants expected *personalized utterances*, while the SFP coach held that these are not necessary for certain types of well-being coaching to work. During Study 2, several participants noted that the robot's emotional expressions were good when they were correct, however they felt they were jarring when the robot assessed their emotions incorrectly. These data indicate that it is difficult to determine the correct level of emotional adaptation, as the robot will inevitably make some mistakes. This is discussed further in Section 6.

5.2.4 Accessibility. Accessibility of a robotic mental well-being coach was seen as both an advantage and disadvantage, particularly in Study 3. In Study 1 and Study 3, participants pointed out that a robotic well-being coach could have more availability than a human coach, with the SFP and Life Coaches in Study 1 remarking that the robot could also be more *neutral* due to lack of daily routines or past life experiences. However, during Study 3, participants mentioned that the cost of a robotic coach such as Pepper could be quite high, and the benefits would have to be high enough to justify that cost. In Study 1, participants remarked that they could see themselves using such a robot if it were easily accessible in a space such as their home or their workplace. However, these options may not be available to all people. Accessibility (as well as the acceptability) of such a robotic coach will depend on the socio-economical and cultural contexts, and should be examined on a case-by-case basis.

Divergences in theme emphasis

5.2.5 Well-being practices and robot roles. Study 1 focused explicitly on what well-being practices a robotic coach could perform, and what it should and shouldn't do in relation to these. Study 2 focused on giving feedback to a robotic coach conducting a *Positive Psychology* session specifically, and Study 3 focused mainly on editing an SFP script. Due to this, *robot roles* were also discussed more in Study 1, since they relate directly to how a robot should conduct itself — i.e., what role it should assume — when delivering well-being practices. During Study 1, well-being practices such as yoga and life-coaching practices were discussed, which weren't discussed during the later studies, which focused on Positive Psychology and SFP specifically. The robot was also seen to be able to for example take the role of a practice buddy during yoga (i.e., doing the practice with the participants), while this robot role did not come up during discussions regarding Positive Psychology and SFP.

5.2.6 Broader advantages. Study 1 was a broader, hypothetical examination of potential robotic coaches for mental well-being. As such, some of the potential advantages of a robotic coach that were expressed in Study 1, were not discussed or seen to be expressed in Study 2 or Study 3, where a robotic application was examined in detail. Advantages that did not receive much discussion in Study 2 or 3 were *neutrality and anonymity* and *reliability and consistency*. *Memory* and *data analysis* were not mentioned in Study 2, which may be due to the one-off nature of the interaction, and these are features that would be visible only in longitudinal interactions. In Study 3, participants noted that they would like to see the robot *recalling previous sessions* and *adapting to the users' emotions*, both previous and current (see Tables 7 and 12).

5.2.7 Phenomena related to HRI. Novelty and habituation effects as well as *suspension of disbelief* were phenomena related to HRI observed in the themes of Study 2 (see Section 4.2.6). As Study 1 did not involve HRI, and Study 3 did not explicitly analyze the HRI itself, these phenomena were not observed in those studies. During the Positive Psychology exercises conducted by the robot Pepper, participants noted that their experience with the robot changed throughout the interaction session,

and that they had to *get used to*, i.e., habituated to, the robot. Participants noted initial feelings of strangeness, and feeling conscious about the experience, which related to both the *novelty* effect as well as the lack of *suspension of disbelief*.

6 DESIGN AND ETHICAL RECOMMENDATIONS

As the result of our qualitative analysis of the three studies described in Sections 4.1, 4.2 and 4.3, we present design guidelines for robotic well-being coaches, selected on the basis of the convergence and divergence of the findings in these studies. The guidelines address the robot's *form*, its *voice*, as well as the role of *user and robot personality*, how the robot should *acknowledge* and *actively listen* to the user, the role of *turn-taking*, the level of *verbal adaptation*, other *adaptation and personalization*. Additionally, we introduce ethical considerations that we think are particularly important when designing robotic coaches for well-being. The guidelines and ethical considerations, collected in Table 14 and 15, were selected to operationalize the results that converged across the three studies, while addressing divergences of results and attempting to resolve these conflicts. These recommendations are given within the context of current robot capabilities and research into robotic well-being coaches. Some recommendations may also be applicable in robots deployed in other contexts, e.g., universities, workplaces, or homes. However, these contexts are out-of-scope for this paper, and researchers should evaluate how each recommendation applies to their robotic coach within their context.

6.1 R1 - Robot form should follow its function

The robot's form, i.e. its appearance and embodiment, should be of a similar level of sophistication to its functionality. In Study 1 (Section 4.1), participants had conflicting opinions on whether they would prefer a humanoid or an abstract robot (see Section 5.2.1). However, participants noted that they would expect the robot's appearance to match its level of communication (i.e. they wouldn't expect a robotic dog to speak, but they would expect conversation abilities from a humanoid robot such as Pepper). As such, we utilised the humanoid robot Pepper in the Positive Psychology exercises (Study 2, Section 4.2) and in the demonstration of Positive Psychology in Study 3 (Section 4.3). Pepper received mainly positive feedback in Study 2, whereas robot form was not discussed in Study 3. Previous works also provide evidence of the relation between the robot's form and user expectations. For example, Haring et al. [51] claimed that when interacting with robots people are biased by visual perception, and they attribute a degree of functionalities to the robot, taking a cognitive shortcut. They showed how this bias can affect the design of a robot and discuss the implications for HRI. Also, Goetz et al. [44] showed that a match between the appearance of the robot and its functionality can improve the collaboration during a human-robot task. In [53], the authors conducted a survey with 56 participants who were asked to assess seven different robotics platforms. Their results showed that participants scored the humanoid robots as most useful.

If designers want to convey conversational capabilities for a robotic well-being coach, we recommend that robots with a more humanoid appearance are chosen. In the case of Mindfulness and Positive Psychology, if these exercises are not administered in a conversational format, abstract or even animal-like robots may be applicable.

6.2 R2 - Robot voice should emphasize variable prosody and slow pace

The use of speech for robots to communicate with humans is a key factor for a successful interaction. While the voice of a robotic well-being coach was not extensively discussed during any of the studies, some recommendations can be made. Pepper's default voice received mainly positive feedback during Study 2 (Section 4.2). However, in Study 3 (Section 4.3), participants noted that a robot's voice may be too monotonous during longer instructional utterances, that introduce the

Table 14. Design and ethical recommendations rationale

Recommendations (HOW)	Issue (WHAT)	Rationale for recommendations (WHY)
Category: Robot Form		
(R1) Robot form should follow its function	The robot's appearance (i.e. humanoid vs. abstract) should match its level of sophistication in its function (e.g. conversational skills).	Participants mentioned this being a requirement in Study 1 (Section 5.2.1). Form and function attribution has been previously examined in HRI design [51]. As different robot forms are matched to different levels of well-being interactions (e.g. humanoid for conversational, pet-like for care-taking), this should be kept in mind.
(R2) Robot voice should emphasize variable prosody and slow pace	Prosody, as well as slow pace, will be particularly important for explanation-heavy (e.g. the Miracle Question in SFP) or instruction-based (e.g. Mindfulness and Meditation) well-being practices.	Participants regarded the robot's voice mainly positively (Section 5.1.2), but during Study 3, participants noted that Pepper's default voice should be slower and use more prosody, in order to engage the user and be understand (Section 4.3). A monotonous voice is more difficult to focus on. Additionally, a slow pace is often used in Meditation to induce calmness [65].
Category: Robot Behaviour (Adaptation and personalization)		
(R3) Acknowledgement and active listening should be implemented through backchanneling and appropriate turn-taking	Backchanneling can be done through body language such as nodding, and phatic expressions. The robot should wait a longer amount of time than in normal conversation for appropriate turn-taking, as users may take more time to reflect on conversational well-being exercises.	Acknowledgement and active listening were important for users to feel listened to (Section 5.1.1). Because verbal adaptation is limited (guideline 1), backchanneling and appropriate turn-taking can address these needs. Backchanneling has been shown to increase perception of active listening [59, 75].
(R4) Verbal adaptation should be limited to preserve well-being practice efficacy	We propose adaptive verbal utterances should be limited to portions of the interaction that do not affect the well-being practice (e.g. introductions, greetings).	Participants call for extensive verbal adaptation in order to feel acknowledged (Section 5.1.1). However, well-being practitioners say extensive verbal adaptation (such as a conversation), could undermine efficacy of well-being practice (Section 5.2.3, [94]). This is a conflict in preferences of coaches and participants. The guideline aims to preserve the efficacy of the well-being intervention and satisfy user requests for verbal adaptation.

Table 15. Design recommendations rationale

Recommendations (HOW)	Issue (WHAT)	Rationale for recommendations (WHY)
Category: Ethical considerations		
(R5) Safeguarding and risk assessment are required prior to HRI	A risk assessment needs to be completed before interaction with the robot, in order to verify that users are not in a mental health crisis. Such users should be directed to other, more appropriate resources.	Participants and coaches noted that a robot would not be appropriate for a person in a mental health crisis (Section 5.1.5). The proposed robotic well-being coach is not to be used with people who need clinical mental health interventions [39, 93].
(R6) Data collection needs informed consent to preserve privacy	Participants should be informed on what data is being collected and why, how it is stored, and how it is used to adapt or personalize the interaction.	Participants call for adaptation and personalization across all studies (Section 5.1.5), but are also concerned about data collection and privacy (Section 5.2.2). Participants note that all users should be informed of data collection, following GDPR [100].
(R7) Researchers should educate users keeping in mind emotional considerations and user characteristics	In addition to educating users about the collection of their data and privacy, users should be informed accurately of the capabilities and limitations of the robot. The education should be given with the content and flexilibity that participants with varying levels of technical literacy can understand it.	User characteristics (e.g., level of experience with technology) may shape the language needed to make user education understand. While the robot being non-judgemental was an advantage across the studies (Section 5.1.3), participants should still be aware of what information they give to the robot (Section 5.1.5), which is especially important in applications with vulnerable users or data [85], and how this data is used. Participants should also be informed that a robot with empathic expression does not actually "possess" empathy as a human does.

well-being exercises to be conducted. A monotonous voice may result very difficult to pay attention to (as discussed specifically in relation to the Miracle Question, which is a method used in SFP).

Within the HRI literature, past works have widely explored the effect of robot's voice in terms of prosody and pace, especially to convey emotions via speech [28, 56, 103]. For instance, Crumpton and Bethel [27] investigated the use of an open source speech synthesizer to convey emotions through the robot's prosody. Their results showed that participants were able to distinguish the emotion conveyed by the robot using its voice. Also, several studies have demonstrated how the robot's voice affect the user perception of the robot's personality [12, 33, 76]. For example, Dou et al. [33] conducted a study with 15 students and a Pepper robot to analyze the correlation between

the perceived personality of the robot the its voice. Their findings showed that the robot with the child-like voice was more likely to be perceived as extroverted, passionate, and relaxed.

We suggest that especially if the robot will be giving long instructions, special attention should be paid to making the robot's voice human-like with varied prosody, and that emphasis is placed on important words in the robot's utterances. This becomes especially important in well-being practices such as Mindfulness and Meditation, where the coach often talks for longer periods of time without input from the coachee, while it conducts guided meditation. During Mindfulness and Meditation sessions, the robotic coach should also have a slow and calm pace of speech, in order to encourage a meditative experience during the practice.

6.3 R3 - Acknowledgement and active listening should be implemented through backchanneling and appropriate turn-taking

Due to the difficulty in creating a sense of the robot *actively listening* (which the participants called for, see Section 5.1.1) through *verbal adaptation*, we suggest that robotic well-being coaches do this through *backchanneling* via *body language* and *phatic expressions*. Phatic expressions (such as "Uh-huh" and "Mm hmm") can be used to convey the active listening of the robot while the user is speaking. Backchanneling has been shown to improve the functioning of human-robot teams [61]. Well-timed movements such as nodding the robot's head when the user is talking have also been previously shown to improve the user's impression of the robot listening to them [78, 84]. Recently, data-driven methods have been used to create believable emotional robotic gestures based on an initial set of hand-designed gestures, in order to introduce a variety of expressions that contributes towards a robot's life-likeness [70]. Such well-timed and believable body language was requested by participants across all three studies.

We suggest that roboticists create robotic well-being coaches carefully designed with appropriately timed backchanneling cues that the robot expresses, especially during conversational well-being practices such as SFP or Positive Psychology.

Our study results show that appropriate timing of robot expressions emerges as being relevant to the smoothness of turn-taking as well. The timing of the robot taking its turn to speak (as expressed via verbal utterances and gestures) should be carefully considered as well. Turn-taking has been previously predicted in human-robot interaction with the observation of backchannels and fillers [50]. Additionally, multimodal perception has been been used to improve the detection of a human's continuing speaking turn by observing their filled pauses and gaze. This significantly reduced robot interruptions, and increased the time people spent speaking to the robot [14]. In Study 2 (Section 4.2), participants commented on how they weren't sure when they should be talking, and when it was the robot's turn. This led to the participants sometimes experiencing the robot interrupting their stream of thought, reducing the experienced efficacy of their self reflection.

We recommend that designers of robotic well-being coaches leave sufficient time between the participant finishing their speaking turn and the robot beginning its next utterance. These times may be longer than would be considered natural during an every-day conversation, due to well-being exercises sometimes requiring the participant to spend more time reflecting on their experience, and thus resulting in longer pauses in the participants' speech. The robot should also be able to discern between the participant being silent and ending their speaking turn, and recognize users using phatic expressions (such as "Hmm") to indicate that they are still continuing their speaking turn after the pause.

6.4 R4 - Verbal adaptation should be limited to preserve well-being practice efficacy

Participants often called for extensive adaptation of the robot's utterances and word variety, and for the robot to personalize its utterances based on what the participant had said (Section 5.1.1).

However, as seen in all three studies, coaches may disagree with this on the basis that it might jeopardize the efficacy of the well-being practice (see Section 5.2.3 for further discussion). In the case of SFP, the structure and the specific phrasing of sessions is carefully structured, and interfering with it could disrupt the clinically-proven efficacy of the sessions. For example, the Miracle Question, a tool used in SFP, uses specific phrasing such as "imagining" a miracle occurring, which encourages coachees to think outside the box [94]. The same concern for preserving efficacy also applies to participants calling for the robot to remind them of what they had said in previous sessions (i.e. creating a memory for the robot) which is not normally done in SFP, as the practice underlines that the coachee should determine important topics to discuss themselves. The coach was of the opinion that the robot could remember what content had been dealt with in previous sessions (i.e. session topics), but not volunteer to the participant what they had previously said, unless they brought it up themselves. In Study 2 (Section 4.2), the coach reviewing the script determined that minimal verbal acknowledgement to the participant's utterances is acceptable, but that the practice has clinical reasoning behind it, and that extensive alterations to approved phrasing may risk efficacy.

In the case of well-being practices based on Mindfulness, the Meditation/Mindfulness coach (study 1, Section 4.1) underlined that the ability to "read the room", and give instructions in the moment as needed, is important. Meditation has previously been flexibly adapted to fit different needs. For example, walking meditation was adapted to older adults by practicing walking meditation at a slower pace, and always with an aid available to support possible loss of balance [74]. This flexibility was also emphasized by the Mindfulness coach in Study 3 (Section 4.3), who noted that the robot should emphasize each participant's specific *strengths*. This ability of the robot to "read the room" and give instructions as needed might need the robot to have sophisticated sensors in order to discern people's level of calmness during individual or group meditation (e.g. from fidgeting, breath rate, etc), and determine when it is appropriate to intervene. Alternatively, this problem could be addressed with a robot intervening at set intervals, even if this does not reach the level of human Mindfulness coach performance. In the case of Mindfulness and Meditation, too frequent interventions could begin to disturb the participant's concentration, thus lessening the efficacy of the practice.

Adaptation is a difficult problem, as users expect a robot, especially a humanoid robot, to adapt to them and *acknowledge* them. However, adapting well-being practices, which often are based on thoroughly tested clinical interventions, can be a difficult and even risky endeavour. The level of adaptation should be based on how appropriate it is for each intervention. Here, we conclude that based on the aforementioned data, SFP should not be extensively adapted, and Mindfulness should be adapted somewhat. As the SFP coach said in Study 3, they would prefer it to have minimal adaptation, to minimize risk of responding incorrectly. This coach noted that the level of improvement in the users' interaction experience by using verbal adaptation may not be worth the risk in the case of failure. Thus, we recommend that roboticists create a "sense" of adaptation by introducing *wording variety* in certain statements that do not affect the procedure of the well-being practices (such as introductions, explaining the content of the sessions, etc.), but not during the well-being content itself.

6.5 R5 - Safeguarding and risk assessment are required prior to interactions

Another concern of participants and coaches across studies was the safeguarding and risk assessment of participants (Section 5.1.5). Recently, HRI literature has been giving more attention to ethical considerations and how it is ethically and morally required to conduct a prior assessment of mental health of users interacting with the robot to safeguard their health [39]. Fiske et al. [39] concluded their paper acknowledging that further research is needed to address the broader ethical and societal concerns of robotic technologies. Also, Stahl and Coeckelbergh [93] discussed ethical aspects, and

they proposed a new framework to evaluate ethical principles and support researchers in this process, namely Responsible Research and Innovation for healthcare robotics.

Analogously, the findings of Study 3 (Section 4.3) suggest that a robot is not equipped to address such situations, and we object to any attempts to do so, as this may pose a danger to the user. While we argue that a robotic mental well-being coach could be useful for non-clinical populations wishing to maintain or improve their well-being, we do not want to see such robots replacing or even stepping in for human well-being professionals in situations where well-being is a matter of clinical intervention. We emphasize the distinction between coaching (focusing on the present and the future, and the flourishing of the coachee) and therapy (more focus on the past, and addressing mental health disorders) [52], where a robotic mental well-being coach can be appropriate for the prior but should not be performing the latter.

Currently, robots do not have adequate capabilities to identify a situation where the user's well-being might be at risk, and as such pre-screening needs to be carried out by a human. If risk to a participants safety arises despite pre-screening, it is important to have a safeguarding procedure established and communicated to the participant beforehand, which will be carried out depending on the context. For example, in the context of a research study, researchers who are not psychologists and not in a clinical relationship with a person will not be able to provide mental health care. Instead, directing participants to appropriate resources with pamphlets and other informational sources can be appropriate (see literature on the impact of pamphlet distribution, e.g., [5, 99]). These limitations of the robot and the researchers should be communicated to the user beforehand: in case they experience discomfort or safety concerns during their interaction with the robot, they can take a pamphlet with them.

In contexts where a robotic coach might be deployed outside of a research setting and it did have the capability to detect safeguarding situations, e.g. a university, a workplace, or at home, appropriate safeguarding needs to be designed on a case-by-case basis. Additionally, the user should be informed and agree to safeguarding procedures beforehand. For example, at home, the user may prefer to use the robot without the robot engaging in safeguarding, as one would use a diary. In other contexts, e.g. at a university, it might be important to alert university healthcare services in potential safeguarding situations. However, such case-by-case recommendations are out of the scope of this paper, and should be researched further in each context, and as robotic safeguarding situation identification capabilities evolve.

We suggest that in the context of current robot capabilities and within research settings, participants be pre-screened by a human before beginning the use of a robotic mental well-being coach, in order to identify situations where a prospective user may be in significant distress or suffering from a mental health crisis.

6.6 R6 - Data collection needs informed consent to preserve privacy

The ethical consideration that received most attention throughout all three studies was data collection and privacy (Section 5.1.5). Previous research has also placed emphasis on the preservation of privacy in digital well-being applications. Privacy control approaches have been detailed with a user-centric approach, especially for well-being applications that collect personalized data with ubiquitous sensors in environments such as the user's home or workplace [18]. Some virtual coaching applications have opted for, e.g., not collecting demographic information to preserve privacy [55].

Our findings show that the consideration for privacy was in conflict with the preference for the robot's *personalization and adaptation*, especially during the group discussion in Study 3 (Section 4.3). Specifically, the Mindfulness coach called for adaptation by the robot to a specific user's strengths, but did not want the robot to collect any personal data. Though efforts were made by

the researchers to explain why both these things could not be achieved, this concept was not necessarily understood by all participants due to lack of technical knowledge. This indicates that while PD strives to elicit design choices from experts and prospective users of robots, these choices may not always be realistic.

In contrast, the SFP coach present in Study 1 (Section 4.1) noted that privacy may be increased when interacting with a robotic coach in comparison to a human coach. The coach noted that during clinical interventions, information about patients is shared with other clinicians (patients are made aware of this). While we do not propose the use of a robotic well-being coach for clinical populations, this presents and interesting case where a robot may in fact preserve privacy, even though it is collecting information to be stored in a more accessible format than a human coach.

We recommend that the data collection and processing performed by a robotic well-being coach should be efficiently communicated to its users, following GDPR. This is especially important as users may have differing levels of technical knowledge, and may not perceive that any information is being recorded unless specifically stated. Users should be informed about the type of data that is being collected, e.g., the content of their utterances, their facial expressions, or their gestures. Users should also be informed and give consent to how this data will be stored and for how long — as is stipulated by GDPR. As discussed in Section 6.5, it should also be explained whether and how participants' data will be used for safeguarding purposes in their particular context. To preserve *transparency* of the robot's operation, users should also be made aware how the robot's behaviour (e.g. selection of utterances) may be affected by the collection of their data.

6.7 R7 - Researchers should educate users keeping in mind emotional considerations and user characteristics

The topic of user education encompasses different concerns that we have identified in our studies, even if not explicitly mentioned by the participants. This ethical consideration aims to address varying *user expectations*, and accurately convey to the user the *capabilities and limitations* of a robotic well-being coach. This involves educating users by giving them an appropriate impression of the robotic coach's capabilities and limitations, especially in the case of users that are less technologically knowledgeable. This is not a simple matter, as emphasizing the robot's limitations too enthusiastically may reduce the effectiveness of well-being exercises being conducted. Especially in applications with vulnerable users, such as people disclosing details about their well-being, should be given accurate information about the operation of the robot due to its emotional proximity [30, 31].

Emotional consideration, i.e., the consideration of how the emotional relating of users to the robot may influence their behaviour [7], is important when designing of robotic well-being coaches. In Study 1 (Section 4.1), the SFP coach noted that when they conduct clinical therapy, they may notice patients developing an *over-reliance* on the therapist, idealizing them or viewing them as the only person that would understand them. While this may not be a big concern in non-clinical well-being coaching, it should still be considered when designing such robotic well-being coaches. Such over-reliance can be mitigated by informing the users that the robot is intended to be used as a tool.

A concern in robot design has been that the design of robot capabilities, such as empathic expressions, can be seen as deceptive or inauthentic [19, 102] However, previous research has also found that expressions of empathy can improve people's experiences with robots [79], and robotic coaches in particular [8?]. We recommend that researchers be aware that while participants express their desire for a robot to have empathic expressions (see Section 5.1.1), and that those expressions can improve the well-being practice, it does introduce some concerns. In order to mitigate these concerns, if the robot uses empathic expressions, we recommend that at the beginning of a study

(or if this is not feasible, at the end of a study), researchers have an in-person discussion with participants. This discussion should inform users that the robot has been designed to express empathy in order to facilitate a better coaching experience [8, 79**?**], but this does not however mean that the robot "possesses" empathy in the human sense. Informing users of robotic capabilities can be helpful in clarifying misconceptions and mitigating feelings of deception [19, 102].

User education is also important in the case of users lacking *technological knowledge* (e.g., elderly users, as suggested by a participant to be a user group to specifically benefit from a robotic coach in Study 2, Section 4.2) to accurately judge the amount of *privacy* that a robotic coach may afford them (as discussed in Section 6.6). Previous studies have noted that especially in the case of vulnerable users or sensitive applications, special attention should be paid to privacy [7]. This concern can be addressed by informing users about their data being collected and processed in comprehensible language, either before beginning any sessions with the robot coach, or prior to each session. We recommend that at the beginning of the sessions, a person knowledgeable about the robot (e.g., the researcher) be available to the participants to answer any questions. The researcher should clarify what the robot is currently capable of and is "good at" (i.e., facilitating non-clinical well-being exercises), and what it is not capable of and is "bad at" (i.e., identifying safeguarding situations or conducting meaningful therapeutic conversations). This way, the participant can form an accurate conception of the robots capabilities and limitations, and can clarify any technical terms or request further explanation from the researcher, if their technological knowledge is a barrier to understanding how the robot works. This especially becomes a concern in the case of commercially available robots, where the informed consent of users may not be as meticulously implemented as in academic studies.

We recommend that designers of robotic well-being coaches take user education into consideration on a user-by-user basis, and provide user education as applicable to their specific application keeping in mind emotional considerations and user characteristics.

7 CONCLUSION, LIMITATIONS AND FUTURE WORK

In this paper, we present the qualitative analysis of three studies focusing on the design and evaluation of robotic mental well-being coaches. Based on this analysis, we present design guidelines including ethical considerations for such robots. Additionally, we discuss what aspects of the design of such a robot should be considered in the future. We present these analyses and recommendations in order to help future designers and roboticists creating robotic well-being coaches, and to operationalize the knowledge we have found from these studies.

We propose the design recommendations and guidelines, as well as the ethical considerations presented in this paper based on the coverage of the three studies, which examined a robotic well-being coach from different perspectives (see Table 13) and achieved data saturation (see Section 3). However, our studies did not examine the longitudinal use of robotic mental well-being coaches (though the motivations to use such a robot in the long-term were discussed in Study 3, Section 4.3). Future research should examine how robots designed with these considerations in mind are experienced by participants in the long term, over multiple interaction sessions. Such examinations of robots could potentially bring forth the need for additional recommendations regarding capabilities that can only be examined in the long term, such as long-term personalization. Additionally, our paper describes an iterative design process to distil recommendations to inform the design and development of robotic coaches for mental well-being limited to an adult target population. The design and ethical recommendations proposed do not address the needs of more vulnerable populations such as children, the elderly, or people with neuro-developmental disorders. Future research should be conducted for designing robotic coaches for such target populations.

There are several aspects that are known to impact HRI we have not discussed in this work and could be addressed in future works. We highlight them as follows.

Cultural Influences. In recent years, researchers have argued the importance of considering different cultures during human-agent interactions [60]; even so, most of the HRI works has not yet incorporated those aspects, designing robots that are still mainly influenced by Western cultures. The few works that focused on HRI cross-cultural research investigated the Western and Eastern cultures represented respectively by the American and Japanese populations [68]. However, even within the same culture (e.g., Western), many differences and nuances are observed (e.g., Italians vs. British people) [26]. In our context, cross-cultural research would be beneficial for creating robotic well-being coaches. Future work could examine how the guidelines presented here are applicable to different cultures and populations, expanding the results of Study 3, with participants remarking that e.g., a user's socio-economical status might affect how they experience the robot.

Role of Context. Lately, the HRI community has highlighted the role of the context during human-robot interactions [22]. In fact, human behaviors, mental processes, and emotional states vary depending on the environment in which they occur [86]. In parallel, many efforts have been made to perceive and model the social and normative context in human-agent interactions within the affective computing field [32, 101]. It is crucial to develop robots that are context-aware to be able to adapt appropriately to various situations during human-robot interactions. Particularly, in our work, designing robotic coaches that are aware of the context could positively impact the well-being practice. Future work could investigate further what is the role of the context in this application scenario.

Cross-fertilization Between Affective Computing and Social Robotics. The affective computing field is advancing in the automated analysis of emotional and social signals. However, social robotics has not yet integrated these most recent developments into robotic platforms [23]. In addition, due to the necessity for real-time processing skills and the lack of computational capacity of the robotic platforms on the market, deploying social robots in real human-robot interaction scenarios is still an open challenge [83]. Future work should focus on the collaboration between the affective computing and social robotics fields by leveraging cloud computing or external sensors as suggested in [83], to design autonomous robots that can deliver well-being exercises.

ACKNOWLEDGMENTS

We would like to thank Dr. Indu Bodala for contributing to Study 1, Nikhil Churamani for contributing to Study 2, and Katherine Parkin for her contributions to Study 1, Study 2 and Study 3. We thank Dr. Milo Phillips-Brown for the insightful conversation on the ethics of robotic coaches. We would like to thank all study participants and coaches for their valuable contributions to this research.

Funding: M. Axelsson is funded by the Osk. Huttunen foundation and the EPSRC under grant EP/T517847/1. M. Spitale and H. Gunes have been supported by the EPSRC/UKRI under grant ref. EP/R030782/1 (ARoEQ). **Open Access:** For open access purposes, the authors have applied a Creative Commons Attribution (CC BY) licence to any Author Accepted Manuscript version arising. **Data Access Statement:** Overall qualitative analysis of research data underpinning this publication is contained in the manuscript, as well as manuscripts related to the original studies as referred to in the manuscript. Additional raw data related to this publication cannot be openly released as the raw data contains and transcripts of the participants' interviews, group discussions, and interaction with the robot, which were impossible to anonymise.

REFERENCES

[1] Nida Abbasi, Micol Spitale, Joanna Anderson, Tamsin Ford, Peter Jones, and Hatice Gunes. 2022. Can Robots Help in the Evaluation of Mental Wellbeing in Children? An Empirical Study. (2022).

[2] Takuto Akiyoshi, Junya Nakanishi, Hiroshi Ishiguro, Hidenobu Sumioka, and Masahiro Shiomi. 2021. A Robot That Encourages Self-Disclosure to Reduce Anger Mood. *IEEE Robotics and Automation Letters* 6, 4 (2021), 7925–7932.

[3] Arion R Alexopoulos, Jake G Hudson, and Oluwatomisin Otenigbagbe. 2020. The use of digital applications and COVID-19. *Community mental health journal* 56, 7 (2020), 1202–1203.

[4] Patrícia Alves-Oliveira, Patrícia Arriaga, Ana Paiva, and Guy Hoffman. 2021. Children as robot designers. In *Proceedings of the 2021 ACM/IEEE International Conference on Human-Robot Interaction*. 399–408.

[5] Eiichi Ashizawa, Midori Yoshioka, Yuko Sunami, and Shinichi Sato. 2018. Use of a pamphlet to encourage individuals to eat slowly to help reduce the development of metabolic syndrome. *Sangyo Eiseigaku Zasshi= Journal of Occupational Health* 61, 1 (2018), 9–15.

[6] Minja Axelsson. 2021. Social Robot Co-Design Canvases. https://doi.org/10.17605/OSF.IO/JG2T8

[7] Minja Axelsson, Indu P Bodala, and Hatice Gunes. 2021. Participatory Design of a Robotic Mental Well-being Coach. In *2021 30th IEEE International Conference on Robot & Human Interactive Communication (RO-MAN)*. IEEE, 1081–1088.

[8] Minja Axelsson, Nikhil Churamani, Atahan Caldir, and Hatice Gunes. 2022. Participant Perceptions of a Robotic Coach Conducting Positive Psychology Exercises: A Systematic Analysis. *arXiv preprint arXiv:2209.03827* (2022).

[9] Minja Axelsson, Raquel Oliveira, Mattia Racca, and Ville Kyrki. 2021. Social Robot Co-Design Canvases: A Participatory Design Framework. *ACM Transactions on Human-Robot Interaction (THRI)* 11, 1 (2021), 1–39.

[10] M. Axelsson, M. Racca, D. Weir, and V. Kyrki. 2019. A Participatory Design Process of a Robotic Tutor of Assistive Sign Language for Children with Autism. In *Proc. RO-MAN 2019*. 1–8. https://doi.org/10.1109/RO-MAN46459.2019.8956309

[11] Minja Axelsson, Micol Spitale, and Hatice Gunes. 2023. Robotic Coaches Delivering Group Mindfulness Practice at a Public Cafe. In *Companion of the 2023 ACM/IEEE International Conference on Human-Robot Interaction*. 86–90.

[12] Matthew P Aylett, Selina Jeanne Sutton, and Yolanda Vazquez-Alvarez. 2019. The right kind of unnatural: designing a robot voice. In *Proceedings of the 1st international conference on conversational user interfaces*. 1–2.

[13] Mandy D Bamber and Joanne Kraenzle Schneider. 2016. Mindfulness-based meditation to decrease stress and anxiety in college students: A narrative synthesis of the research. *Educational Research Review* 18 (2016), 1–32.

[14] Miriam Bilac, Marine Chamoux, and Angelica Lim. 2017. Gaze and filled pause detection for smooth human-robot conversations. In *2017 IEEE-RAS 17th International Conference on Humanoid Robotics (Humanoids)*. IEEE, 297–304.

[15] Holly Blake, Fiona Bermingham, Graham Johnson, and Andrew Tabner. 2020. Mitigating the psychological impact of COVID-19 on healthcare workers: a digital learning package. *International journal of environmental research and public health* 17, 9 (2020), 2997.

[16] Júlia Pareto Boada, Begoña Román Maestre, and Carme Torras Genís. 2021. The ethical issues of social assistive robotics: A critical literature review. *Technology in Society* 67 (2021), 101726.

[17] Indu P. Bodala, Nikhil Churamani, and Hatice Gunes. 2021. Teleoperated Robot Coaching for Mindfulness Training: A Longitudinal Study. In *2021 30th IEEE International Conference on Robot Human Interactive Communication (RO-MAN)*. 939–944. https://doi.org/10.1109/RO-MAN50785.2021.9515371

[18] Wouter Bokhove, Bob Hulsebosch, B Van Schoonhoven, Maya Sappelli, and Kees Wouters. 2012. User privacy in applications for well-being and well-working: requirements and approaches for user controlled privacy. (2012).

[19] Marion Boulicault, Milo Phillips-Brown, Jacqueline M Kory-Westlund, Stephanie Nguyen, et al. 2021. Authenticity and Co-design: On Responsibly Creating Relational Robots for Children. *Algorithmic Rights and Protections for Children* (2021).

[20] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative research in psychology* 3, 2 (2006), 77–101.

[21] Bengisu Cagiltay, Hui-Ru Ho, Joseph E Michaelis, and Bilge Mutlu. 2020. Investigating Family Perceptions and Design Preferences for an In-Home Robot. In *Proc. IDC 2020* (London, United Kingdom) *(IDC '20)*. Association for Computing Machinery, 229–242.

[22] David Cameron, J Aitken, E Collins, Luke Boorman, Adriel Chua, Samuel Fernando, Owen McAree, Uriel Martinez Hernandez, and James Law. 2015. Framing factors: The importance of context and the individual in understanding trust in human-robot interaction. (2015).

[23] Oya Celiktutan, Evangelos Sariyanidi, and Hatice Gunes. 2018. Computational analysis of affect, personality, and engagement in human–robot interactions. In *Computer Vision for Assistive Healthcare*. Elsevier, 283–318.

[24] Jiaee Cheong, Micol Spitale, and Hatice Gunes. 2023. "It's not Fair!" – Fairness for a Small Dataset of Multi-Modal Dyadic Mental Well-being Coaching. In *2023 11th International Conference on Affective Computing and Intelligent Interaction (ACII)*. IEEE.

[25] Nikhil Churamani, Minja Axelsson, Atahan Caldir, and Hatice Gunes. 2022. Continual Learning for Affective Robotics: A Proof of Concept for Wellbeing. In *2022 10th International Conference on Affective Computing and*

Intelligent Interaction Workshops and Demos (ACIIW). IEEE.
[26] Daniela Conti, Allegra Cattani, Santo Di Nuovo, and Alessandro Di Nuovo. 2015. A cross-cultural study of acceptance and use of robotics by future psychology practitioners. In *2015 24th IEEE international symposium on robot and human interactive communication (RO-MAN)*. IEEE, 555–560.
[27] Joe Crumpton and Cindy L Bethel. 2015. Validation of vocal prosody modifications to communicate emotion in robot speech. In *2015 International Conference on Collaboration Technologies and Systems (CTS)*. IEEE, 39–46.
[28] Joe Crumpton and Cindy L Bethel. 2016. A survey of using vocal prosody to convey emotion in robot speech. *International Journal of Social Robotics* 8, 2 (2016), 271–285.
[29] Lúzie Fofonka Cunha, Lucia Campos Pellanda, and Caroline Tozzi Reppold. 2019. Positive psychology and gratitude interventions: A randomized clinical trial. *Frontiers in psychology* 10 (2019), 584.
[30] Kate Darling. 2015. 'Who's Johnny?'Anthropomorphic framing in human-robot interaction, integration, and policy. *Anthropomorphic Framing in Human-Robot Interaction, Integration, and Policy (March 23, 2015). ROBOT ETHICS* 2 (2015).
[31] Kate Darling. 2016. Extending legal protection to social robots: The effects of anthropomorphism, empathy, and violent behavior towards robotic objects. In *Robot law*. Edward Elgar Publishing.
[32] Giorgio De Magistris, Riccardo Caprari, Giulia Castro, Samuele Russo, Luca Iocchi, Daniele Nardi, and Christian Napoli. 2022. Vision-Based Holistic Scene Understanding for Context-Aware Human-Robot Interaction. In *International Conference of the Italian Association for Artificial Intelligence*. Springer, 310–325.
[33] Xiao Dou, Chih-Fu Wu, Kai-Chieh Lin, and Tzu-Min Tseng. 2019. The effects of robot voice and gesture types on the perceived robot personalities. In *International Conference on Human-Computer Interaction*. Springer, 299–309.
[34] Yunfei Duan, Myung Yoon, Zhixuan Liang, and Johan Ferdinand Hoorn. 2021. Self-Disclosure to a Robot: Only for Those Who Suffer the Most. *Robotics* 10, 3 (2021), 98.
[35] Brian R Duffy and Karolina Zawieska. 2012. Suspension of disbelief in social robotics. In *2012 IEEE RO-MAN: The 21st IEEE International Symposium on Robot and Human Interactive Communication*. IEEE, 484–489.
[36] Juliane Eberth and Peter Sedlmeier. 2012. The effects of mindfulness meditation: a meta-analysis. *Mindfulness* 3, 3 (2012), 174–189.
[37] Robert A Emmons and Charles M Shelton. 2002. Gratitude and the science of positive psychology. *Handbook of positive psychology* 18 (2002), 459–471.
[38] Noelle Fields, Ling Xu, Julienne Greer, and Erin Murphy. 2021. Shall I compare thee… to a robot? An exploratory pilot study using participatory arts and social robotics to improve psychological well-being in later life. *Aging & mental health* 25, 3 (2021), 575–584.
[39] Amelia Fiske, Peter Henningsen, Alena Buyx, et al. 2019. Your robot therapist will see you now: ethical implications of embodied artificial intelligence in psychiatry, psychology, and psychotherapy. *Journal of medical Internet research* 21, 5 (2019), e13216.
[40] Kathleen Kara Fitzpatrick, Alison Darcy, and Molly Vierhile. 2017. Delivering cognitive behavior therapy to young adults with symptoms of depression and anxiety using a fully automated conversational agent (Woebot): a randomized controlled trial. *JMIR mental health* 4, 2 (2017), e7785.
[41] Jayde AM Flett, Harlene Hayne, Benjamin C Riordan, Laura M Thompson, and Tamlin S Conner. 2019. Mobile mindfulness meditation: a randomised controlled trial of the effect of two popular apps on mental health. *Mindfulness* 10, 5 (2019), 863–876.
[42] Fabian Gander, René T Proyer, and Willibald Ruch. 2016. Positive psychology interventions addressing pleasure, engagement, meaning, positive relationships, and accomplishment increase well-being and ameliorate depressive symptoms: A randomized, placebo-controlled online study. *Frontiers in psychology* 7 (2016), 686.
[43] Domenico Giacco, Neelam Laxhman, and Stefan Priebe. 2018. Prevalence of and risk factors for mental disorders in refugees. In *Seminars in cell & developmental biology*, Vol. 77. Elsevier, 144–152.
[44] Jennifer Goetz, Sara Kiesler, and Aaron Powers. 2003. Matching robot appearance and behavior to tasks to improve human-robot cooperation. In *The 12th IEEE International Workshop on Robot and Human Interactive Communication, 2003. Proceedings. ROMAN 2003*. Ieee, 55–60.
[45] Lewis R Goldberg. 1993. The structure of phenotypic personality traits. *American psychologist* 48, 1 (1993), 26.
[46] Anthony M Grant and M Cavanagh. 2010. Life coaching. *The complete handbook of coaching* (2010), 297–310.
[47] Anthony M Grant and Michael J Cavanagh. 2014. The solution-focused approach to coaching. *The complete handbook of coaching* (2014), 327–344.
[48] L Suzzy Green, Lindsay G Oades, and Anthony M Grant. 2006. Cognitive-behavioral, solution-focused life coaching: Enhancing goal striving, well-being, and hope. *The Journal of Positive Psychology* 1, 3 (2006), 142–149.
[49] Greg Guest, Arwen Bunce, and Laura Johnson. 2006. How many interviews are enough? An experiment with data saturation and variability. *Field methods* 18, 1 (2006), 59–82.

[50] Kohei Hara, Koji Inoue, Katsuya Takanashi, and Tatsuya Kawahara. 2018. Prediction of turn-taking using multitask learning with prediction of backchannels and fillers. *Listener* 162 (2018), 364.

[51] Kerstin S Haring, Katsumi Watanabe, Mari Velonaki, Chad C Tossell, and Victor Finomore. 2018. FFAB—The form function attribution bias in human–robot interaction. *IEEE Transactions on Cognitive and Developmental Systems* 10, 4 (2018), 843–851.

[52] Vicki Hart, John Blattner, and Staci Leipsic. 2001. Coaching versus therapy: A perspective. *Consulting Psychology Journal: Practice and Research* 53, 4 (2001), 229.

[53] S Maryam Fakhr Hosseini, Samantha Hilliger, Jaclyn Barnes, Myounghoon Jeon, Chung Hyuk Park, and Ayanna M Howard. 2017. Love at first sight: Mere exposure to robot appearance leaves impressions similar to interactions with physical robots. In *2017 26th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*. IEEE, 615–620.

[54] Annika Howells, Itai Ivtzan, and Francisco Jose Eiroa-Orosa. 2016. Putting the 'app'in happiness: a randomised controlled trial of a smartphone-based mindfulness intervention to enhance wellbeing. *Journal of happiness studies* 17, 1 (2016), 163–185.

[55] Alexis M Jeannotte, Derek M Hutchinson, and Gabriella R Kellerman. 2021. Time to Change for Mental Health and Well-being via Virtual Professional Coaching: Longitudinal Observational Study. *Journal of medical Internet research* 23, 7 (2021), e27774.

[56] Eun-Sook Jee, Yong-Jeon Jeong, Chong Hui Kim, and Hisato Kobayashi. 2010. Sound design for emotion and intention expression of socially interactive robots. *Intelligent Service Robotics* 3, 3 (2010), 199–206.

[57] Sooyeon Jeong, Sharifa Alghowinem, Laura Aymerich-Franch, Kika Arias, Agata Lapedriza, Rosalind Picard, Hae Won Park, and Cynthia Breazeal. 2020. A Robotic Positive Psychology Coach to Improve College Students' Wellbeing. In *Proc. RO-MAN 2020*. 187–194.

[58] Karen R Johnson, Sunyoung Park, and Sanghamitra Chaudhuri. 2020. Mindfulness training in the workplace: Exploring its scope and outcomes. *European Journal of Training and Development* (2020).

[59] Susanne M Jones. 2011. Supportive listening. *The Intl. Journal of Listening* 25, 1-2 (2011), 85–103.

[60] Michiel P Joosse, Ronald W Poppe, Manja Lohse, and Vanessa Evers. 2014. Cultural differences in how an engagement-seeking robot should approach a group of people. In *Proceedings of the 5th ACM international conference on Collaboration across boundaries: culture, distance & technology*. 121–130.

[61] Malte F Jung, Jin Joo Lee, Nick DePalma, Sigurdur O Adalgeirsson, Pamela J Hinds, and Cynthia Breazeal. 2013. Engaging robots: easing complex human-robot teamwork using backchanneling. In *Proceedings of the 2013 conference on Computer supported cooperative work*. 1555–1566.

[62] Raida Karim, Yufei Zhang, Patrícia Alves-Oliveira, Elin A Björling, and Maya Cakmak. 2022. Community-Based Data Visualization for Mental Well-being with a Social Robot.. In *HRI*. 839–843.

[63] Ariane Kerst, Jürgen Zielasek, and Wolfgang Gaebel. 2020. Smartphone applications for depression: a systematic literature review and a survey of health care professionals' attitudes towards their use in clinical practice. *European archives of psychiatry and clinical neuroscience* 270, 2 (2020), 139–152.

[64] M Kim, K Oh, J Choi, J Jung, and Y Kim. 2011. User-centered HRI: HRI research methodology for designers. In *Mixed Reality and Human-Robot Interaction*. Springer, 13–33.

[65] Glenn E Knowlton and Kevin T Larkin. 2006. The influence of voice volume, pitch, and speech rate on progressive relaxation training: application of methods from speech pathology and audiology. *Applied psychophysiology and biofeedback* 31, 2 (2006), 173–185.

[66] James Laurie and Ann Blandford. 2016. Making time for mindfulness. *International journal of medical informatics* 96 (2016), 38–50.

[67] Hee Rin Lee, Selma Šabanović, Wan-Ling Chang, Shinichi Nagata, Jennifer Piatt, Casey Bennett, and David Hakken. 2017. Steps toward participatory design of social robots: mutual learning with older adults with depression. In *Proc. HRI 2017*. 244–253.

[68] Velvetina Lim, Maki Rooksby, and Emily S Cross. 2021. Social robots on a global stage: establishing a role for culture during human–robot interaction. *International Journal of Social Robotics* 13, 6 (2021), 1307–1333.

[69] Shane J Lopez, Jennifer Teramoto Pedrotti, and Charles Richard Snyder. 2018. *Positive psychology: The scientific and practical explorations of human strengths*. Sage publications.

[70] Mina Marmpena, Fernando Garcia, Angelica Lim, Nikolas Hemion, and Thomas Wennekers. 2022. Data-driven emotional body language generation for social robotics. https://doi.org/10.48550/ARXIV.2205.00763

[71] John McLeod. 2001. Using grounded theory. *Qualitative research in counselling and psychotherapy* (2001), 71–89.

[72] Sanika Moharana, Alejandro E Panduro, Hee Rin Lee, and Laurel D Riek. 2019. Robots for joy, robots for sorrow: community based robot design for dementia caregivers. In *Proc. HRI 2019*. IEEE, 458–467.

[73] Elaine Mordoch, Angela Osterreicher, Lorna Guse, Kerstin Roger, and Genevieve Thompson. 2013. Use of social commitment robots in the care of elderly people with dementia: A literature review. *Maturitas* 74, 1 (2013), 14–20.

[74] Natalia E Morone and Carol M Greco. 2014. Adapting mindfulness meditation for the older adult. *Mindfulness* 5, 5 (2014), 610–612.

[75] Nasim Motalebi, Eugene Cho, S Shyam Sundar, and Saeed Abdullah. 2019. Can Alexa be your therapist? How back-channeling transforms smart-speakers to be active listeners. In *Conference Companion Publication of the 2019 on Computer Supported Cooperative Work and Social Computing*. 309–313.

[76] Yi Mou, Changqian Shi, Tianyu Shen, and Kun Xu. 2020. A systematic review of the personality of robot: Mapping its conceptualization, operationalization, contextualization and effects. *International Journal of Human–Computer Interaction* 36, 6 (2020), 591–605.

[77] Alissa J Mrazek, Michael D Mrazek, Casey M Cherolini, Jonathan N Cloughesy, David J Cynman, Lefeba J Gougis, Alex P Landry, Jordan V Reese, and Jonathan W Schooler. 2019. The future of mindfulness training is digital, and the future is now. *Current Opinion in Psychology* 28 (2019), 81–86.

[78] Michael Murray, Nick Walker, Amal Nanavati, Patricia Alves-Oliveira, Nikita Filippov, Allison Sauppé, Bilge Mutlu, and Maya Cakmak. 2022. Learning Backchanneling Behaviors for a Social Robot via Data Augmentation from Human-Human Conversations. In *Conference on Robot Learning*. PMLR, 513–525.

[79] Andreea Niculescu, Betsy van Dijk, Anton Nijholt, Haizhou Li, and Swee Lan See. 2013. Making social robots more attractive: the effects of voice pitch, humor and empathy. *International journal of social robotics* 5 (2013), 171–191.

[80] Warren T Norman. 1963. Toward an adequate taxonomy of personality attributes: Replicated factor structure in peer nomination personality ratings. *The journal of abnormal and social psychology* 66, 6 (1963), 574.

[81] World Health Organization. 2022. *World Health Organization - Mental Health issues*. https://www.who.int/health-topics/mental-health

[82] Zola Marie Packman. 1991. The Incredible and the Incredulous: The Vocabulary of Disbelief in Herodotus, Thucydides, and Xenophon. *Hermes* 119, H. 4 (1991), 399–414.

[83] Chung Hyuk Park, Raquel Ros, Sonya S Kwak, Chien-Ming Huang, and Séverin Lemaignan. 2020. Towards Real World Impacts: Design, Development, and Deployment of Social Robots in the Wild. , 600830 pages.

[84] Hae Won Park, Mirko Gelsomini, Jin Joo Lee, and Cynthia Breazeal. 2017. Telling stories to robots: The effect of backchanneling on a child's storytelling. In *2017 12th ACM/IEEE International Conference on Human-Robot Interaction (HRI*. IEEE, 100–108.

[85] Laurel Riek and Don Howard. 2014. A code of ethics for the human-robot interaction profession. *Proceedings of we robot* (2014).

[86] Hanan Salam and Mohamed Chetouani. 2015. A multi-level context-based modeling of engagement in human-robot interaction. In *2015 11th IEEE international conference and workshops on automatic face and gesture recognition (FG)*, Vol. 3. IEEE, 1–6.

[87] Brian Scassellati, Henny Admoni, and Maja Matarić. 2012. Robots for use in autism research. *Annual review of biomedical engineering* 14 (2012), 275–294.

[88] Martin EP Seligman. 2007. Coaching and positive psychology. *Australian Psychologist* 42, 4 (2007), 266–267.

[89] Elena Sheldon, Melanie Simmonds-Buckley, Claire Bone, Thomas Mascarenhas, Natalie Chan, Megan Wincott, Hannah Gleeson, Karmen Sow, Daniel Hind, and Michael Barkham. 2021. Prevalence and risk factors for mental health problems in university undergraduate students: A systematic review with meta-analysis. *Journal of Affective Disorders* 287 (2021), 282–292.

[90] Clay Spinuzzi. 2005. The methodology of participatory design. *Technical communication* 52, 2 (2005), 163–174.

[91] Micol Spitale, Minja Axelsson, and Hatice Gunes. 2023. Robotic Mental Well-Being Coaches for the Workplace: An In-the-Wild Study on Form. In *Proceedings of the 2023 ACM/IEEE International Conference on Human-Robot Interaction* (Stockholm, Sweden) *(HRI '23)*. Association for Computing Machinery, New York, NY, USA, 301–310. https://doi.org/10.1145/3568162.3577003

[92] Micol Spitale and Hatice Gunes. 2022. Affective Robotics For Wellbeing: A Scoping Review. In *2022 10th International Conference on Affective Computing and Intelligent Interaction Workshops and Demos (ACIIW)*. IEEE, 1–8.

[93] Bernd Carsten Stahl and Mark Coeckelbergh. 2016. Ethics of healthcare robotics: Towards responsible research and innovation. *Robotics and Autonomous Systems* 86 (2016), 152–161.

[94] Elijah Steinbrecher, Sara Smock Jordan, and Brie Turns. 2020. Providing immediate hope to survivors of natural disasters: A miracle question intervention. *The American Journal of Family Therapy* 49, 2 (2020), 204–219.

[95] Ladislav Timulak. 2009. Meta-analysis of qualitative studies: A tool for reviewing qualitative research findings in psychotherapy. *Psychotherapy Research* 19, 4-5 (2009), 591–600.

[96] Ladislav Timulak. 2014. Qualitative meta-analysis. *The SAGE handbook of qualitative data analysis* 481 (2014).

[97] John Torous, Jessica Lipschitz, Michelle Ng, and Joseph Firth. 2020. Dropout rates in clinical trials of smartphone apps for depressive symptoms: a systematic review and meta-analysis. *Journal of affective disorders* 263 (2020), 413–419.

[98] European Union. 2022. *Complete guide to GDPR compliance*. https://gdpr.eu/

[99] John T Vetto, Polly M Dubois, and Irene Perez Vetto. 1996. The impact of distribution of a patient-education pamphlet in a multidisciplinary breast clinic. *Journal of Cancer Education* 11, 3 (1996), 148–152.

[100] Paul Voigt and Axel Von dem Bussche. 2017. The eu general data protection regulation (gdpr). *A Practical Guide, 1st Ed., Cham: Springer International Publishing* 10, 3152676 (2017), 10–5555.

[101] Peng Wang, Hongyi Liu, Lihui Wang, and Robert X Gao. 2018. Deep learning-based human motion recognition for predictive context-aware human-robot collaboration. *CIRP annals* 67, 1 (2018), 17–20.

[102] J Kory Westlund, Cynthia Breazeal, and A Story. 2015. Deception, secrets, children, and robots: What's acceptable. In *Workshop on The Emerging Policy and Ethics of Human-Robot Interaction, held in conjunction with the 10th ACM/IEEE International Conference on Human-Robot Interaction*.

[103] Katie Winkle and Paul Bremner. 2017. Investigating the real world impact of emotion portrayal through robot voice and motion. In *2017 26th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*. IEEE, 627–634.

[104] Katie Winkle, Praminda Caleb-Solly, Ailie Turton, and Paul Bremner. 2018. Social robots for engagement in rehabilitative therapies: Design implications from a study with therapists. In *2018 13th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 289–297.

[105] Katie Winkle, Praminda Caleb-Solly, Ailie Turton, and Paul Bremner. 2019. Mutual shaping in the design of socially assistive robots: A case study on social robots for therapy. *International Journal of Social Robotics* (2019), 1–20.

A LIST OF CODES

Robotic well-being coach codebook	Study 1	Study 2	Study 3	Total
Robot capabilities	**1**	**1**	**1**	**3**
Active listening		1		1
Acknowledgement		1		1
Adaptation		1	1	2
Adaptation to emotions			1	1
Contextual adaptation			1	1
Personalized utterances			1	1
Automateable therapy and meditation elements	1			1
Customization			1	1
Engagement	1			1
Expressing empathy	1		1	2
Instructing and demonstrating	1			1
Feedback and evaluation of practice	1			1
Instructing and demonstrating	1			1
Interactive and adaptive responses and conversation		1		1
Keywords		1		1
Pattern analysis and progress tracking	1			1
Personalization	1		1	2
Reminders to practice	1			1
Responsiveness		1		1
Safety, e.g. safeguarding	1		1	2
Turn-taking				1
Wording variety		1		1
Robot features	**1**	**1**	**1**	**3**
Appearance / Form	1	1		2
Behaviour	1	1		2
Environment	1			1

Interaction	1			1
Expressivity	1			1
Movement, gestures		1		1
Personality, social skills			1	1
Voice, prosody		1		1
Robot advantages	**1**	**1**	**1**	**3**
Accessibility	1	1	1	3
Data analysis	1			1
Embodiment			1	1
Lack of judgement		1	1	2
Memory	1			1
Neutrality and anonymity	1			1
Presence	1	1		2
Reliability and consistenty	1			1
Saying out loud		1		1
Robot disadvantages	**1**	**1**	**1**	3
Accessibility			1	1
Barriers of cultural background			1	1
Incorrect responses		1		1
Invasive data collection	1			1
Lack of humanness	1	1		2
Lack of understanding			1	1
Privacy issues	1	1		2
Safety issues	1			1
Technical limitations			1	1
Technology malfunction or other problems	1			1
Timing		1		1
Unwanted robot behaviour	1			1

Ethical considerations			**1**	**1**
Data collection and privacy			1	1
Safeguarding and risk assesment		1	1	2
VS human or application		**1**		**1**
VS human or application		1		1
VS application		1		1
Motivation to start using robot	**1**			**1**
Accessibility	1			1
Adaptive and responsive	1			1
Community of users	1			1
Ease of use, low maintenance	1			1
Recommendations	1			1
Motivation to keep using robot / long-term motivation		**1**	**1**	**2**
Adaptive and responsive		1		1
Ease of use, low maintenance		1		1
Number of sessions			1	1
Privacy respected		1		1
Recall of previous sessions			1	1
Seeing results		1		1
Users / User expectations	**1**	**1**		**2**
Attitudes	1			1
Evidence-based change in opinion	1			1

Neutrality	1		1
Openness or optimism	1		1
Skepticism	1		1
Covid-19		1	1
Introverts		1	1
Level of interactivity (coach vs aid)	1		1
Nurseries and hospitals		1	1
Robot form influencing expectations	1		1
Robot roles	**1**		**1**
Companionship	1		1
Instructor	1		1
Peer or friend for wellbeing practice	1		1
Taking care of robot (robot as pet)	1		1
Telepresence	1		1
Working alongside human coach	1		1
Robot-led wellbeing practice	**1**		**1**
Coaching practice	1		1
Emotion logging	1		1
Meditation	1		1
Mindfulness	1		1
Other	1		1
SFP	1		1
Yoga	1		1
PP exercises		**1**	**1**
Correct questions		1	1
Difficulty understanding tasks		1	1

Experience of helpfulness	1		1
Instructions to use robot	1		1
Thinking of new things	1		1
Repeat questions	1		1
Phenomena		**1**	**1**
Novelty or habituation		1	1
Suspension of disbelief		1	1
Conditions	**1**		**1**

B STUDY 1: PROSPECTIVE USER INTERVIEW

An example prospective user interview. We modified interviews according to the participant's level of active well-being practice: currently doing well-being practices, considering doing well-being practices, and previously did well-being practices but stopped. Additionally, we tailored interview questions based on participants' responses to a survey prior to the interview. This example for the semi-structured interview is for a participant who is currently doing well-being practices.

(1) Current well-being practices
- You said you are currently doing well-being practices. Could you say what makes you motivated to do them?
- Are you experiencing particular benefits? In your survey you specify reduced anxiety, improved performance, better sleep, developed gratitude, increased happiness, and reduced stress. Could you elaborate on this?
- Are there particular challenges you are addressing with well-being practices?

(2) Questions about specific survey answers
- In the survey you say you are currently doing mindfulness and meditation, what made you decide to do these things? Had you seen or heard of them before? In your survey you specify personal interest and research.
- You said you are aiming to achieve reduced anxiety and improved performance and better sleep and develop gratitude and increase happiness and reduce stress with the practices. Could you elaborate?
- You said that you are practicing once a week. What motivates you to keep practicing at this frequency?
- You've practiced since *[year omitted]*. What made you stick with it?
- In your survey you specify having a routine to your practice. Could you elaborate on that?

(3) Tools used for well-being practice
- You say you are using non-digital tools: note taking by hand and in-person practice with an instructor. Could you elaborate on how you're using them, and when?
- What benefits are you finding with these tools?
- Are these benefits motivating you to practice, and to stick with the practice?
- You say you were previously using digital tools. Could you tell me a bit about them?

(4) Technology providing help
- You mention that you stopped using technology during your practice because it was inconvenient and you were not satisfied with the quality of the technology and it felt impersonal. How did you experience these things?
- Did you experience any benefits when using technology?
- Did you experience or could you imagine experiencing any benefits to using technology for your practice?
- Are there any additional features that would make you want to use technology for your practice again?
- Is there something that would make you stop using technology again?
- Did you have any concerns about using technology for well-being practice?
- Have you seen or heard about some interesting well-being practices employing technology in some way?
- Do you have ideas about how technology could be used to create well-being practices?
- Does something else come to your mind about technology and well-being practices?

C STUDY 1: COACH INTERVIEW

An example semi-structured coach interview structure, here for the Life Coach. All interviews shared a general structure, and were then tailored specifically to the coach. Some questions were based on background information the coach had given prior to the interview. Prior to questions about robots, all coaches were shown the same five photos and videos of robots used with the prospective users in the group discussions.

(1) Types of practices
- What types of coaching practices do you instruct?
- Do you do one-to-one or group instructions? If both, what are the differences?
- Could you briefly expand on the concept of Life Coaching?
- You mentioned you use positive psychology as part of coaching, would you please expand on that? How do you use PS in the context of life coaching?
- You mentioned that you use cognitive behaviour psychology as part of coaching, would you please expand on that? How do you use cognitive behaviour psychology in the context of life coaching?

(2) Benefits and goals
- What do you think are the benefits of life coaching for the coachee ?
- What have your coachees told you they think are the benefits?
- What goals do the coachees usually have?
- How are these goals identified?

(3) Coachees' motivations
- What types of needs do people have when they come for coaching?
- How do the coachees find out about your services and life coaching? Do they come through recommendations / word of mouth?

(4) Practicalities
- How long are coaching sessions usually?
- How often are coaching sessions usually?
- How many sessions are there usually?

(5) Tools and technology
- Do you use tools when you instruct / coach? What kinds of tools?
- How are the tools helping the coachee?
- How are the tools helping you as the coach?
- Do you use technology as a tool when you coach? What kinds of technology?
- How is technology helping the coachee?
- How is technology helping you as the coach?
- What advantages and disadvantages do you see practicing online virtually, vs. face-to-face?
- What are the differences when the sessions are mediated through technology vs. having face-to-face sessions?
- Do you think there could be benefits to using technology during COVID-19?

(6) Further technology
- How do you think technology could help further with coaching?
- Are you aware of technologies that have helped people with their coaching or their goals as coachees? (e.g. Headspace)
- Have you ever used such technologies?
- What do you think the benefits of such technologies are?

- Do you think such technologies can become the primary method of delivery for coaching? Or rather, are they secondary and should be used in support with a human instructor?

(7) Specific questions about practice
- How do you decide which goals to work on with clients?
- Who is in charge of the coaching sessions you instruct? Who leads the interaction?
- How do you decide when to talk and when to stay silent?
- Do you use a script during the sessions?
- How do you adapt the sessions? According to the day? According to the person?

(8) Robots
- Would you be able to imagine such a robot providing or giving a life coaching session?
- Which aspects of life coaching do you think would be appropriate in this context, and why?
- Are there any other coaching practices that you think would be appropriate to be delivered by such robots? Why?

(9) Robot advantages and disadvantages
- What do you think the advantages of using a robot for these types of practices may be? Previous participants have mentioned for example: personalization and adaptation, analyzing previous practice patterns and progress, accessibility (having it in your house/work), consistency, and being non-judgemental.
- How about the disadvantages? Previous participants have mentioned for example: Privacy, expensiveness, repetitive interactions, unreliable technology, and intrusive reminders.

(10) Robot capabilities
- Some participants noted they like self-disclosure done by their coach — do you use self-disclosure (i.e. sharing your own experiences) as part of coaching instruction?
- Do you think you express emotions during sessions? How?
- Do people need guidance during the coaching sessions?
- Do people get sidetracked when exploring an area or a particular topic?
- Do people need to be given some time to think?

D STUDY 2: POST-INTERACTION INTERVIEW

The questions asked during the semi-structured post-interaction interview:

(1) Robot
- What do you think was good about the robot's appearance? What did you like?
- What do you think was bad about the robot's appearance? What would you change?
- How appropriate do you think the robot was for helping you focus on the positive aspects in your life? How appropriate do you think the robot was for helping you foster a positive attitude toward things that happen in your life?
- How appropriate do you think the robot was for these coaching tasks?
- How useful do you think the robot was for these coaching tasks?
- How beneficial do you think the robot was for these coaching tasks?

(2) Behaviour
- What do you think was good about the robot's behaviour? What did you like?
- What do you think was bad about the robot's behaviour? What would you change?
- How appropriate do you think the robot's behaviour was for these coaching tasks?
- How useful do you think the robot's behaviour was for these coaching tasks?

- How beneficial do you think the robot's behaviour was for these coaching tasks?
- Do you think the robot understood what you said?
- Do you think the robot understood how you felt?
- Do you think the root adapted to what you said and did?
- Did you notice any differences?

E SFP SCRIPT

' *This is the first session of BSFP (Brief-Solution Focused Practice), in a series of at least 1-4 sessions. Please comment and make edits on the dialogue of the robot according to what you would like it to say. The example given by the human participant is imaginary.*

1. Introduction

Robot: Hello, My name is X. What should I call you?

Human: NAME

Robot: Nice to meet you NAME

Robot: So we've got 4 sessions together, each 20 minutes long, with about a week between each session. We're going to be using a Brief Solution-Focused approach. It's a type of conversation that helps people think about what their best hope and goals for the future are, and think about what strengths and resources they have as a person. And how they might know that they are getting closer to those best hopes, and noticing signs of progress towards these best hopes, just through conversation. Noticing what works and doing more of it.

Robot: And what we find is that when we talk about people's best hopes and signs of progress, that they naturally start noticing what works well, and doing more of what works well, without really trying. That's kind of the approach. Do you have any questions about that? Please say 'yes' or 'no'.

Human: No.

{Human: Yes.

Robot: I will call in the experimenter to answer your questions. Experimenter, could you please come in?}

Session begins

2. What are your best hopes / what are you hoping to achieve

Robot: So thinking about what your best hopes are, how would you know that these sessions have been useful to you? What would you like to work on?

Human: I would like to have a better mood, I often feel sleepy.

Robot: So, you would like to have a better mood and feel less sleepy.

Human: Yes

3. Miracle question

Robot: Say you leave here today, and you do what you have to do for the rest of the evening, and you go to sleep. And something odd happens in the night, so during the night a miracle happens. And your best hopes are realized, that's what the miracle is. And you wake up the next morning, but you were asleep when the miracle happened, so you don't know that it happened. What would be the first signs that would tell that a miracle had happened? You don't know it did. What would you notice?

Human: I would feel excited about my day, and be able to exercise in the morning. I wouldn't feel miserable. I would get things done for the day, and follow my routine.

Robot: So you'd wake up, you'd feel excited, you wouldn't lie in, you would get going. You would follow your routine, what is your routine?

Human: I would go for a run in the morning, and go to work early. And I would do something I enjoy after work. And actually set time aside for doing it, and not feel tired. Have a nice dinner and go to sleep early.

Robot: Are there days where it felt like what you're describing?

4. Are there instances that are good, how do you notice?

Robot: Are there instances when you notice you have achieved your best hopes, in the past or the recent past? When you've woken up and done all the things you've said?

Human: Sometimes yes.

Robot: What percentage of days would you say that you achieve your best hopes?

Human: Maybe 5 days a week.

Robot: What happens on the days or instances where you achieve your best hopes? What leads to you achieving your best hopes on those days?

Human: I get up early, and I eat well.

Robot: So on the days when you achieve your best hopes, you get up early and you eat well?

Human: Yes and I notice I feel better on those days.

Robot: That makes sense.

Robot: Anything else that makes the days when you achieve your best hopes go better?

Human: If I see friends, I notice I feel better that day.

Robot: So you may have already noticed some signs of moving towards your best hopes. That often happens before people come to a session. They've already started working on the problem they're describing.

5. Scaling question

Robot: So where would you put yourself on the scale of 1-10 on the bad days? So 10 would be the miracle has happened/your best hopes have been realised and 1 would be 'you couldn't be further from this'.

Human: 2

Robot: And what about good days?

Human: 5

Robot: So you manage to sometimes get yourself from a 2 to a 5.

Human: Yes, on the better days.

Robot: So there are already things you are doing that can put you up on the scale.

Human: Yes.

6. First signs that things are shifting

Robot: What would be the first signs that things were starting to shift toward your best hopes?

Human: I would notice more of these days in the month.

Robot: Any other small things that you might notice?

Human: I think I would notice that I read my book more, since that's what I like to do when I'm not too tired.

7. Summary

Robot: So in summary, you said you would/wouldn't notice…
You would have more time…
Some signs this would already be happening are…
There has already been a change in the last few weeks where...

8. End of session

Robot: So we are coming up to the 20 minute mark. So my suggestion, every week will be very similar, so I will always suggest that you spend the week noticing signs, any small signs, that you're getting closer to your best hopes. So you don't need to do anything differently, just notice times when you feel you are achieving or feeling your best hopes/.

Robot: And it can be 30 seconds when you feel that way, or a whole week when you feel like that was a great week, like you feel at a 7 on the scale. It doesn't really matter the time, but it's just tuning in to those small moments. And I emphasize small signs, because people expect there to be something big, when actually it's the small things that kind of accumulate. A small sign can even be planning to do something according to your best hopes, even if you don't do it.

Robot: So next week I will ask you what has been better in relation to your best hopes. And if you keep an eye out over the next week, it makes it easier. Sometimes the best hopes can also be maintaining things as is for now. Does that make sense?

Human: yes.

Robot: The nice thing about solution-focused approach is that you're probably already doing things, you're just not noticing them. We're all doing things that work, but we're not paying attention. We tend to notice when things go wrong, but sometimes it can take a bit longer to notice when things go right. In solution-focused practice, we have a structured way to think through things, and having a conversation about it often helps. And the ideas might start to naturally come.

Robot: Thank you for the session.

Human: Thank you!

The previous text was the potential first session. Now please imagine that you are interacting with the robotic coach long-term, over at least 4 sessions. Please edit and comment the following section accordingly.

9. Long-term interaction – Example of Week 2

Scaling question

Robot: So last time we talked about the miracle question, and how you would know that a miracle had happened over night. So what do you think, on the scale of 0 to 10, so 10 is that the miracle happened, and 0 is couldn't be further from that. So where would you put yourself on this scale this week?

Human: 6

Robot: And what makes you a 6 rather than a 5. How do you know you're a 6?

Human: I'm feeling less tired this week, I notice I can wake up earlier and I feel energized enough to exercise in the morning.

Robot: And is there anyone who would notice it first, if we think of someone who knows you really well, do you have an example in mind?

Human: My housemates have seen me around the house earlier in the morning.

Robot: Who do you think of the people who know you, would be the first to notice that you were maybe a 7 on the scale?

Human: I think my housemate, because they also wake up very early in the morning and we meet each other in the kitchen when I'm feeling more energized.

Robot: And what would you notice that told you you were a 7 on the scale?

Human: I think I would be even more energized in the mornings, and maybe prepare a better breakfast than the sandwich I'm having now. Maybe some eggs.

Robot: When you are able to do the things you want, how does that affect your mood?

Human: I feel I'm doing better and I feel more energized because of the better breakfast.