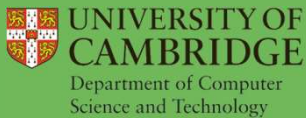


# Robotic Coaches for Mental Wellbeing: From the Lab to the Real World

**Hatice Gunes**

Full Professor of Affective Intelligence & Robotics

University of Cambridge, Department of Computer Science & Technology



@HatijeGyunesh

@AFAR\_Cambridge

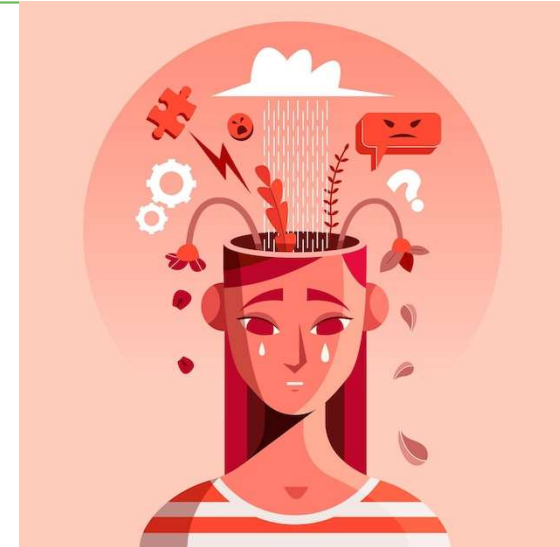
<https://www.cl.cam.ac.uk/~hg410>

<https://cambridge-afar.github.io>

# Motivation

---

- World Health Organization (WHO) reports that mental health conditions have increased +13% in the last decade
- Problem
  - gap between those who require care and those who have access
- **Potential solution**
  - robots can help assess and promote mental wellbeing by offering affordable and accessible practices and services



# Wellbeing Coaching

---

- Coaching goals are to increase the coachee's:
  - hope
  - goal-striving
  - general well-being
- Different styles of coaching:
  - Cognitive behavioural
  - Positive psychology
  - Mindfulness



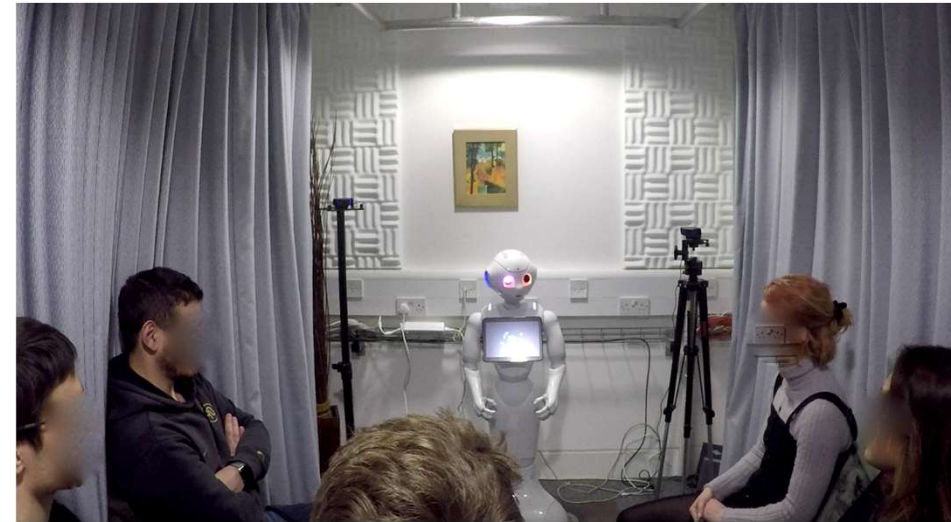
# Social HRI Research Landscape

<b>Study setting</b>	<b>Robot autonomy</b>	<b>Robot form</b>	<b>Study length &amp; frequency</b>	<b>Involvement</b>
In the lab	Wizard of Ozz	Mechanical	One-off	Human coach or therapist
In the wild	Teleoperated	Toy-like	Multi-session	Potential users
...	Pre-scripted	Zoomorphic	...	Human coach & potential users
...	Autonomous	Humanoid	...	...
...	Adaptive	...	...	...
<b>Real-world deployment</b>	<b>Personalized</b>	<b>Comparative</b>	<b>Longitudinal</b>	<b>User-centred and iterative</b>



# Social Robotics for Mental Wellbeing @ Cambridge AFAR Lab

- **Goal:** Be available where humans cannot be and intervene before issues are exacerbated
- **Our Vision**
  - Autonomous wellbeing coach
  - Embodied multimodal interactions
  - Long-term, personalized HRI
- **Our Approach**
  - Iterative design approach
    - Learning from experienced human coaches
    - Face-to-face studies to gather interaction data and design requirements.



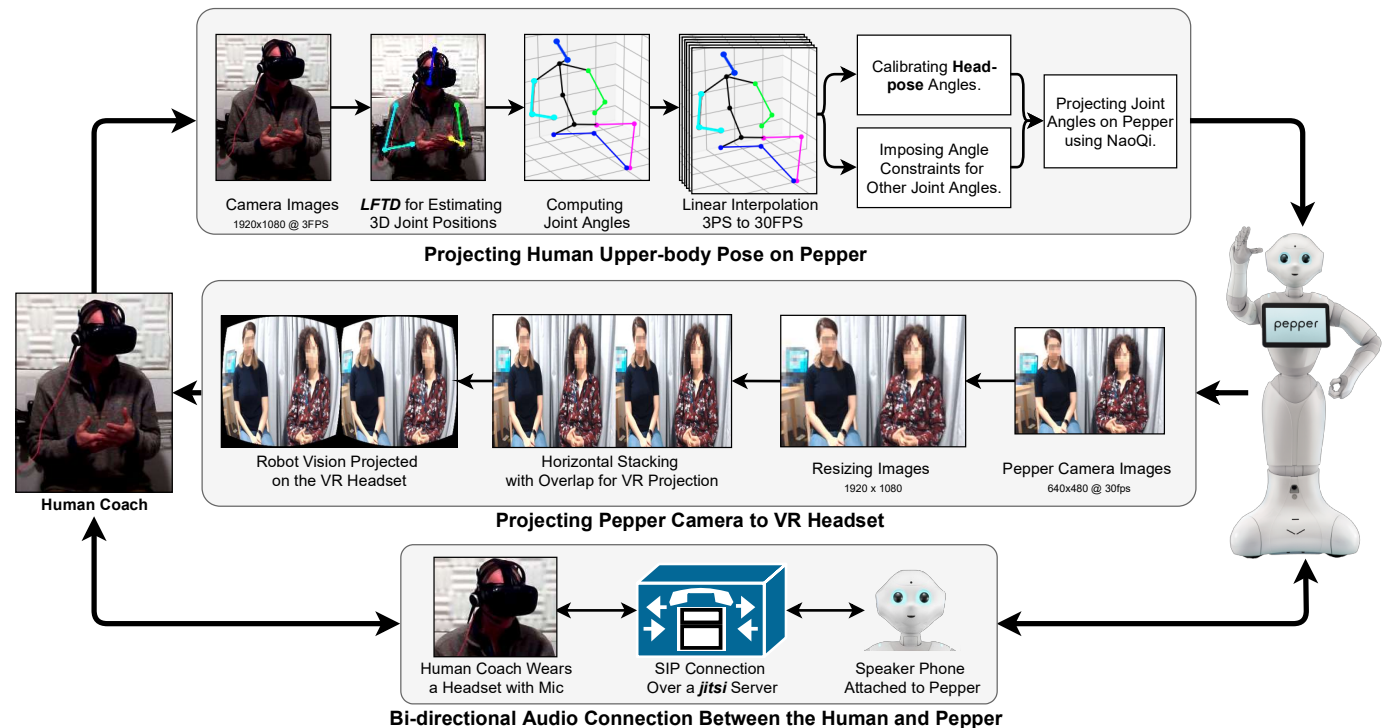
Mindfulness Session delivered by Pepper Robot

# Teleoperated Robot Coaching for Mindfulness Training

## Experiment design

- 5-week mindfulness training delivered by – human coach (**HC**) & a teleoperated robot coach (**RC**)

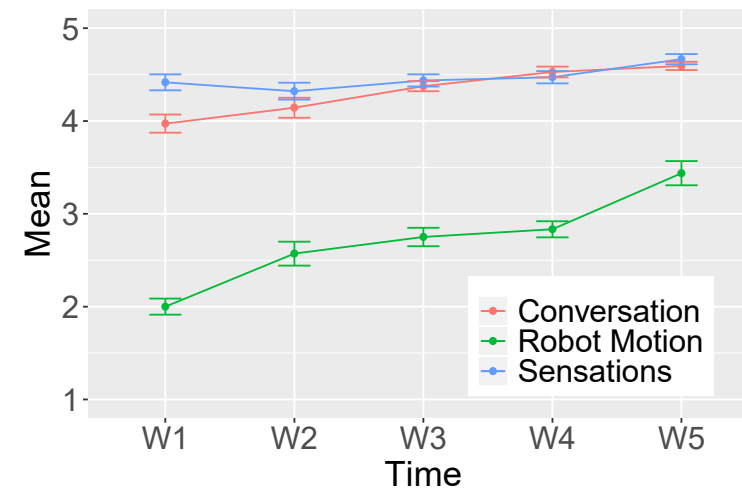
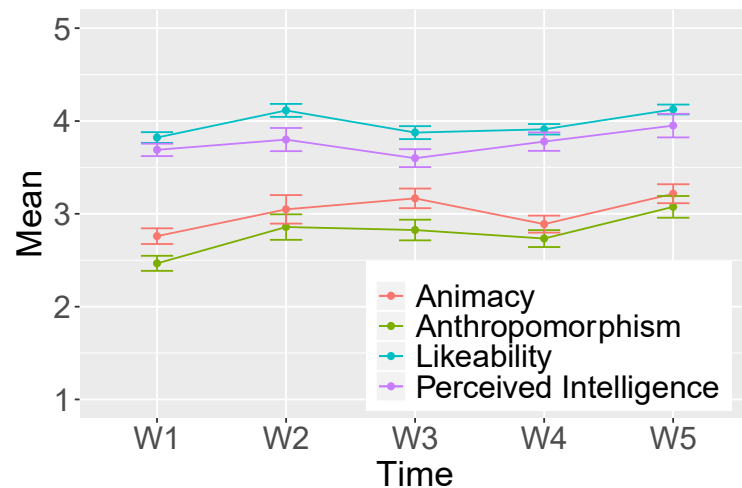
*Pose, vision and audio pipelines during robot teleoperation*



# Longitudinal Changes in the Session Experience Ratings

## ○ Longitudinal interactions with RC

- Significant increase in the *Robot Motion* and *Conversation* ratings with time

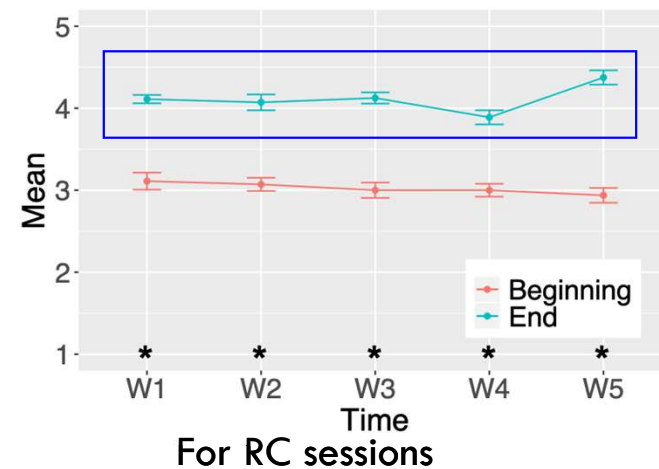
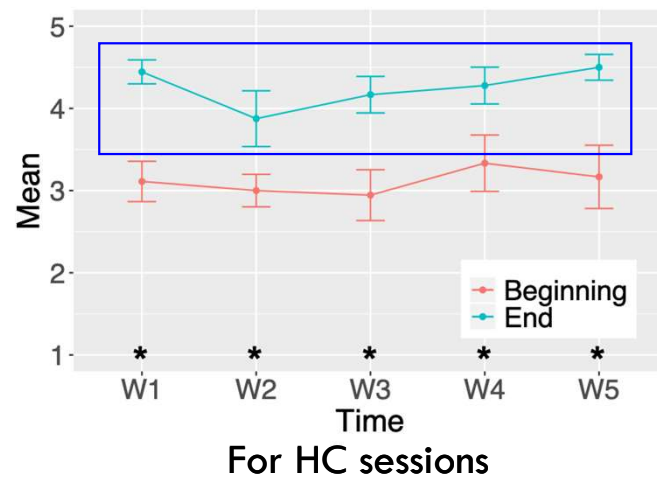


Univariate random-intercept modelling with *time* as within-subject factor.

# Longitudinal Changes in the Session Experience Ratings

## ○ Comparing Feelings at the beginning vs. at the end

- Each session promoted significantly positive mood in the participants for both HC and RC



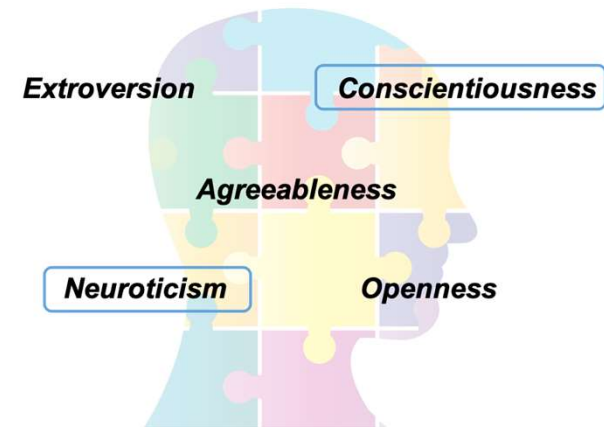
Weeks with significant differences are indicated by \*



# Findings & Limitations Informing the Next Study

- **In-the-lab** robotic coaching study
- Robotic coach is **not autonomous**
- **High conscientious** people expect the robot to move more naturally
- People **high** along **neuroticism** enjoyed the robot sessions less
  - Will robotic wellbeing coaching benefit from **personalization** and **adaptation** ?
- What else the potential **users want**?
- What **other wellbeing practices** could the robotic coach deliver?

## Big Five Personality Traits



# Participatory Design: Data Gathering

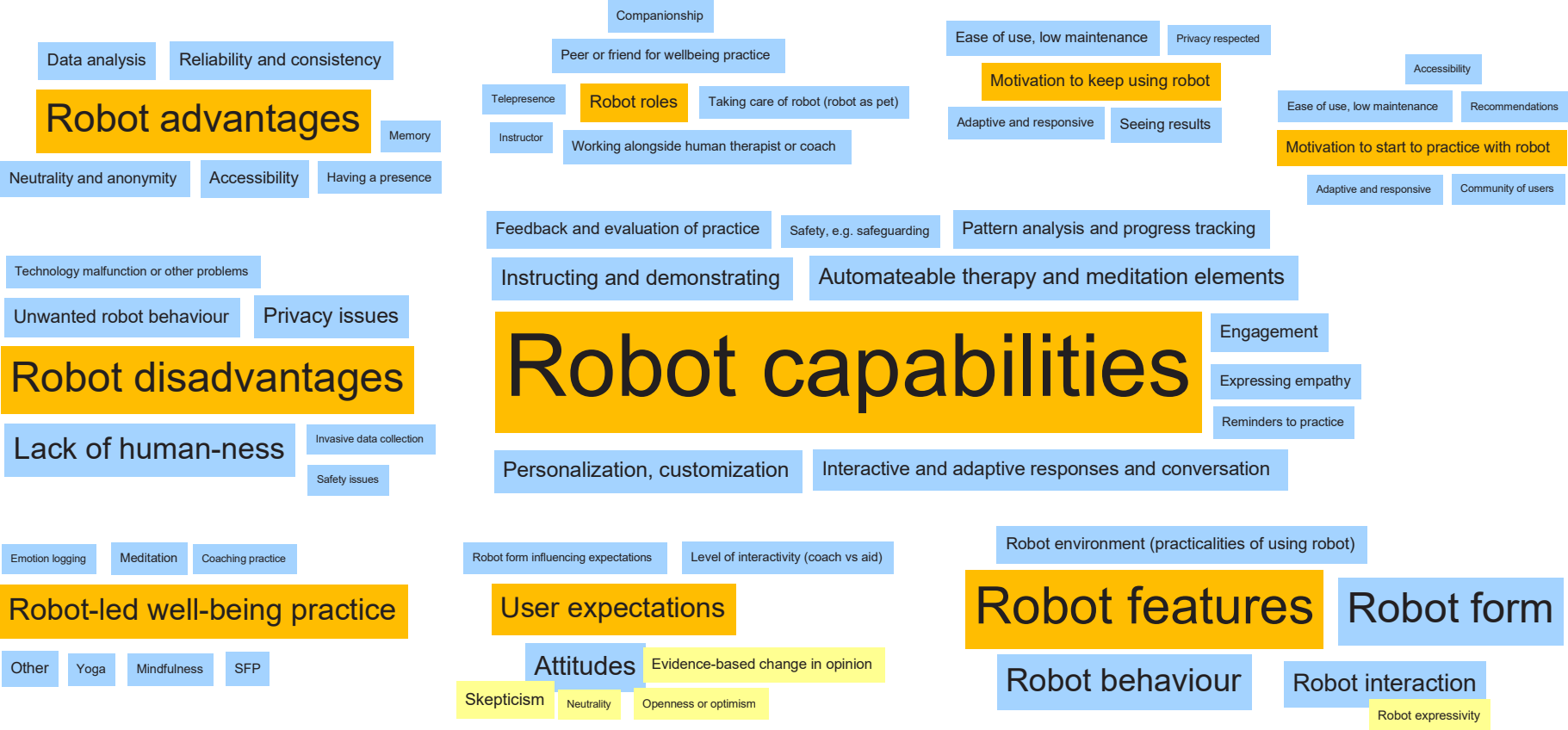
- 8 prospective users
- 3 well-being coaches
  - Mindfulness / Meditation
  - Solution-Focused Practice
  - Life Coaching
- Interviews & focus group discussions
- **Rich qualitative data**

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

# Results

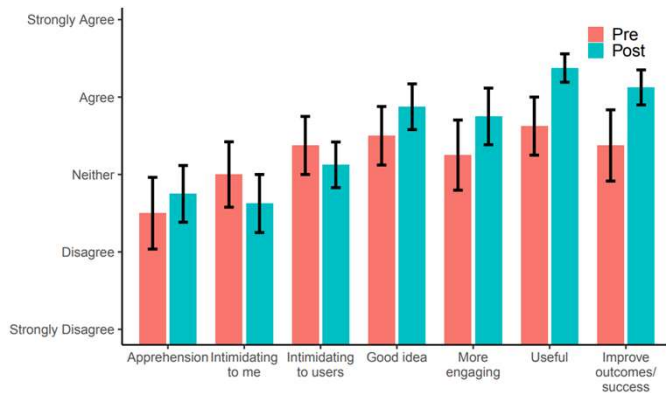


# Thematic Analysis: Results



## Findings Informing the Next Study

- Participants receptive to scientific evidence, more **open to using robot**
- Coaches thought **robot could perform** certain **well-being practices**

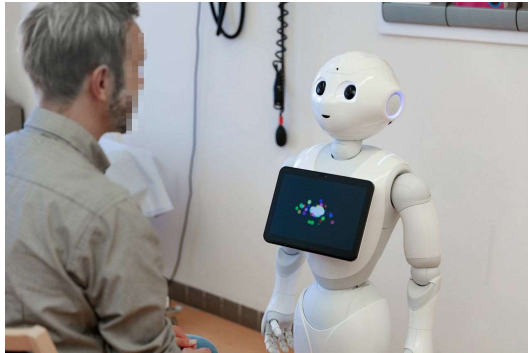


Mean and standard errors of participant ratings on attitudes towards a robotic coach pre and post focus group.

P5: “I am always receptive to evidence, if it has been shown to be beneficial I would certainly give it a try, it lessens my skepticism.”

C3: “... [the robot] could give the person the sense [that] someone is there for you, present.”

# Towards Autonomous & Adaptive Robotic Wellbeing Coach



(a)



(b)



(c)

Interact and adapt.

Extend learning with  
other users.

Adapt to different user  
demographics.



# Traditional vs. Continual Learning

## Traditional

- Models **trained in isolation** on benchmark datasets.
- Large datasets enable **generalisation** across contexts.
- Training data** might be very **different** from **application** scenarios.
- Generalisation comes at the cost of learning individual differences.
- Cumbersome to** retrain and **update** models.

## Continual Learning

- Agents acquire and **integrate knowledge incrementally** about changing environments.
- Data** only made **available sequentially**.
- Highly sensitive towards **changing data conditions**.
- Adaptations in learning to **avoid forgetting**.

### CL Problem Formulation:

$$A_i^{CL} : \langle h_{i-1}, Tr_i, M_{i-1}, t_i \rangle \rightarrow \langle h_i, M_i \rangle$$

New Data
Task
Updated Experience

Model
Experience
Improved Model

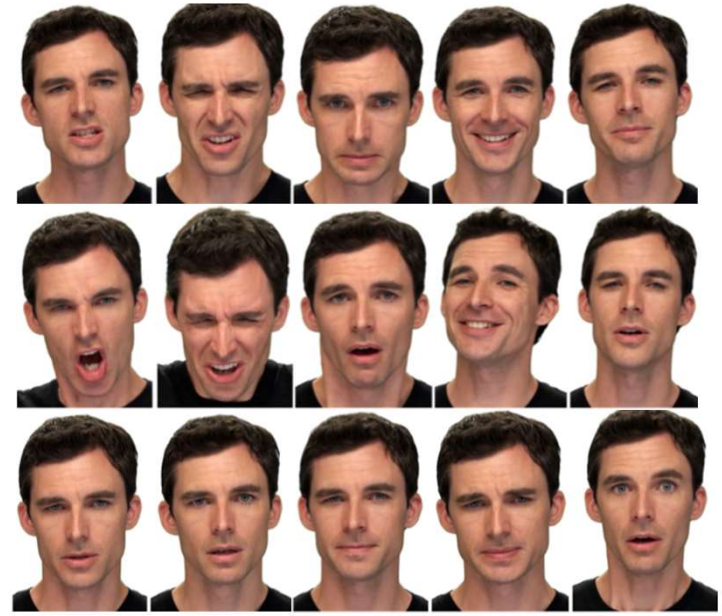
# Traditional vs. Continual Learning

**Generalisation** for facial expression recognition



(a)

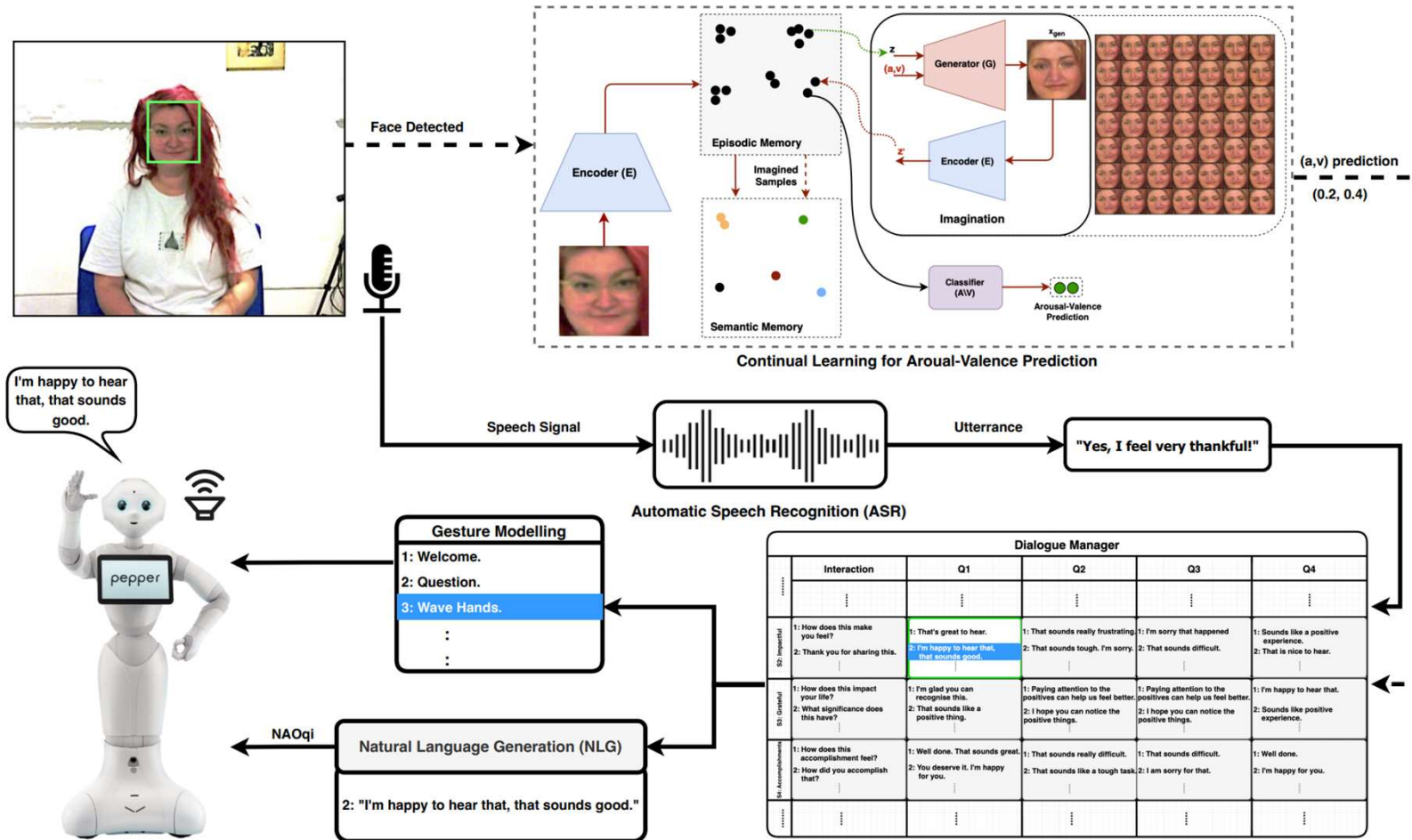
**Personalisation** to learn individual expressions



(b)



# Continual Learning for Affective and Wellbeing Robotics

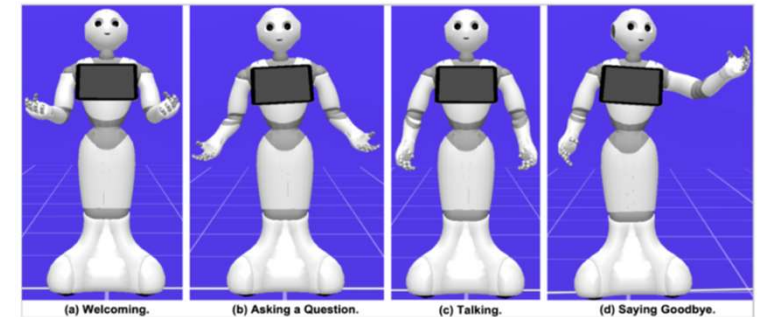


# Continual Learning in Action: Wellbeing Coaching

- Positive psychology exercises
  - 2 impactful things in their lives in the recent past
  - 2 things that they felt grateful for in the recent past
  - 2 recent accomplishments in the recent past
  
- Conditions
  - C1 - Static and Scripted Interaction
  - C2 - Affect-based Adaptation without Personalisation
  - C3 - Affect-based Adaptation with Continual Personalisation

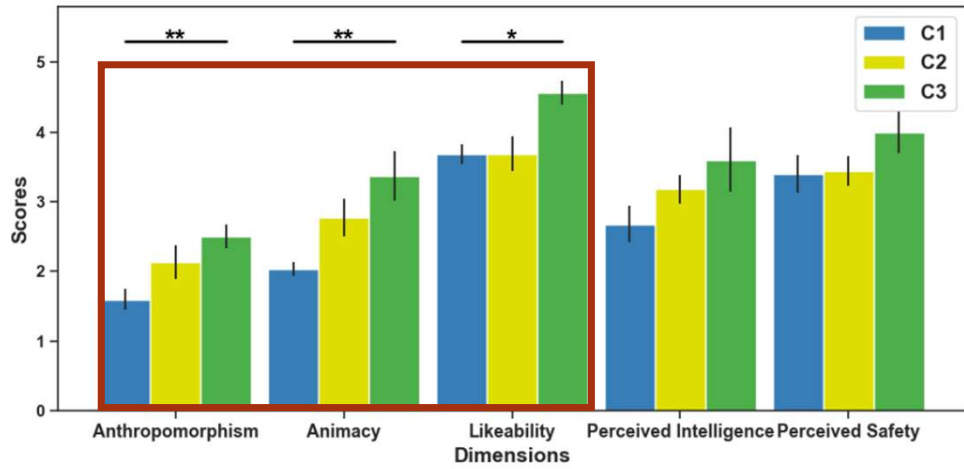


Setup: Pepper interacting with the Participant.

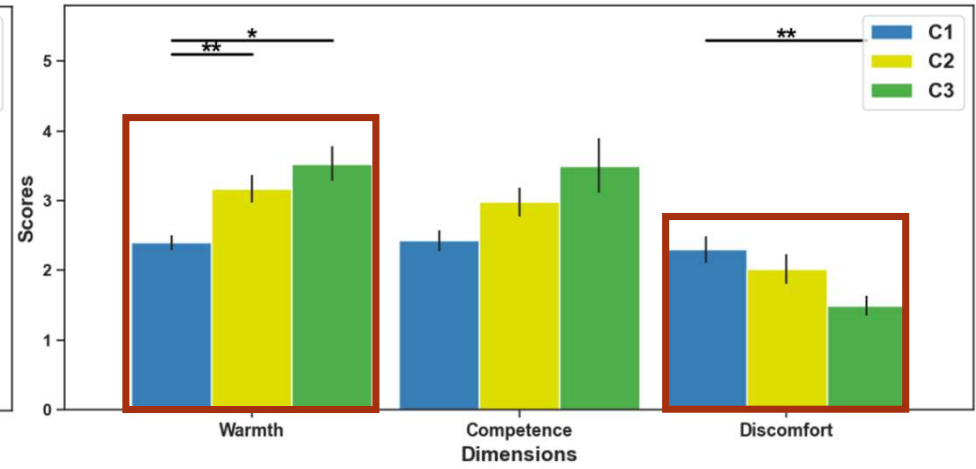


Pepper displaying gestures during the interactions.

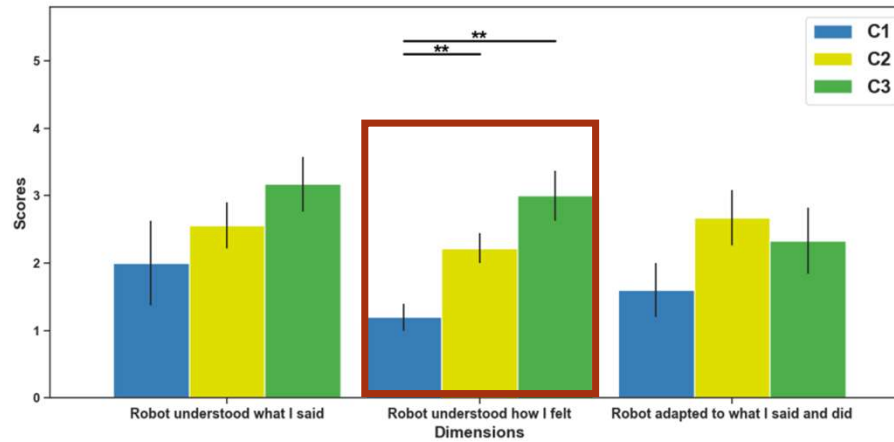
# Results



(a) GODSPEED Scores.



(b) RoSAS Scores.



Customised Question Scores under C1, C2, and C3 conditions.

Scores for C1, C2, and C3 conditions.  
 \* represents  $p < 0.05$  and \*\* represents  $p < 0.01$ .

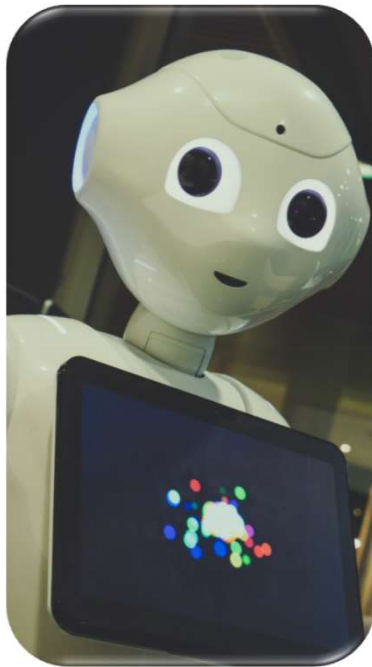
# Findings & Limitations Informing the Next Study

---

- **In-the-lab** robotic coaching
- Robotic coach **lacks** in **speech** capabilities (understanding / generation)
- A **unimodal** affect **perception** model be **insufficient**
  - explore multimodal perception to better asses participants' (affective) responses
- Continual Learning in action comes with **challenges**



# Continual Learning: Challenges & Recommendations



RECOMMENDATIONS FOR AFFECTIVE ROBOTICS

Recommendation	Why is this important and needed?	How can this be achieved?
Acquire person-specific data	Adapting learning models to individual preferences requires large amounts of data that can only be sourced through interactions with users.	(1) Conduct introductory HRI rounds to enable the robot to collect additional data about the user. (2) Leverage adversarial learning to train a generative model to simulate additional person-specific data.
Obtain normative baselines	The robot needs to know the behavioural <i>norm</i> for each user against which deviations can be observed. Deviations help identify shifts in user socio-emotional behaviours and infer changes in interaction context.	(1) Conduct interactions under contextually inert (neutral) situations during introduction rounds. (2) Use the (subtle) deviations from this baseline, given the interaction context, to analyse shifts.
Extract semantic associations	Adapting the learning for a large number of users is computationally intractable. Learning models will get saturated, not able to remember previous information or learn with new individuals.	(1) Form user groupings, using person-specific attributes ( $C_u$ in Eq. 2-3) to learn group-based adaptations. (2) Use unsupervised data clustering to facilitate learning semantic groupings of users.
Learn contextual affordances	Interactions are driven by context and humans switch between contexts without clear boundaries. Contextual attributions may not always be implicit and need to be learnt separately	(1) Learn context-aware embeddings to distinguish between task boundaries. (2) Use contextual affordances (e.g. $T_i$ in Eq. 3) to facilitate smooth switching between affective HRI contexts.
Balance memory with computation	The memory-computation trade-off needs to be considered w.r.t the application domain. Adding more memory facilitates rehearsal of past knowledge, while additional computation power improves adaptation to novel experiences.	(1) Use generative models for pseudo-rehearsal to reduce model's memory foot-print. (2) Offload part of the computation/memory load to RaaS-based solutions to balance old vs. novel learning.
Allow controlled forgetting	When learning is continuous, redundant information in the memory/model, is not released, hindering learning capacity of the model.	(1) Utilise forgetting mechanisms (inspired by biological organisms) on unused memory locations or parts of the model, to learn new knowledge.
Use multiple performance metrics	Benchmark evaluations from conventional ML and CL perspectives are needed for reproducibility and fairness guarantees, and to evaluate model's robustness to dynamic shifts in data distributions.	(1) Report CL performance metrics (Section II-E), along with the classification metrics of F-measure and AUC-ROC scores or reward-function dynamics for behaviour learning.

# Findings & Limitations Informing the Next Study

---

- **In-the-lab** robotic coaching study
- Robotic coach **lacks** in **speech** capabilities (understanding / generation)
- A **unimodal** affect **perception** model be **insufficient**
  - explore multimodal perception to better asses participants' (affective) responses
- Continual Learning in action comes with **challenges**
- Investigate **longitudinal interactions** over time
  - to determine whether the **effects hold** (long-term HRI)

# Robotic Mental Well-being Coaches in the Workplace

- Research questions



# Robotic Mental Well-being Coaches in the Workplace

- Research questions

## RQ1

How does the robot form influence coachees' perceptions of the robotic coach in the workplace?





# Robotic Mental Well-being Coaches in the Workplace

- Research questions



## RQ2

How do employees perceive the robotic coach's personality, and do the perceptions differ due to form?



# Robotic Mental Well-being Coaches in the Workplace

- Research questions



**RQ3**  
How do the perceptions of the coachee-coach alliance (working alliance) differ across the two forms?

# Robotic Mental Well-being Coaches in the Workplace



QT - QT Robot



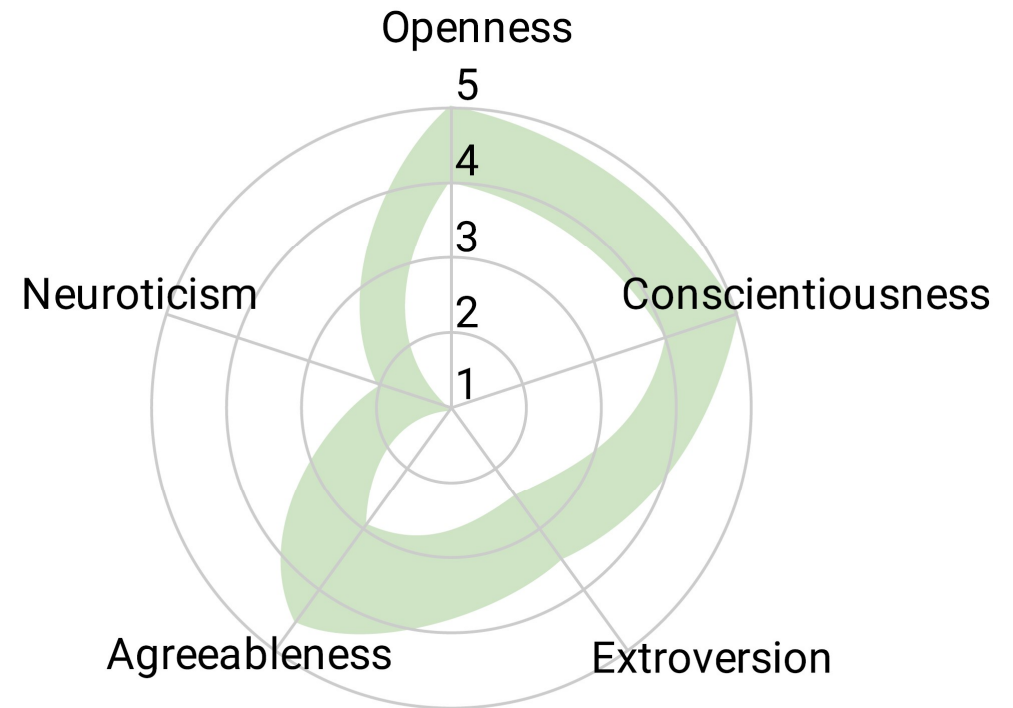
M - Misty Robot

# Robotic Mental Well-being Coaches in the Workplace

## ○ Robot Personality

Robotic mental well-being coach personality:

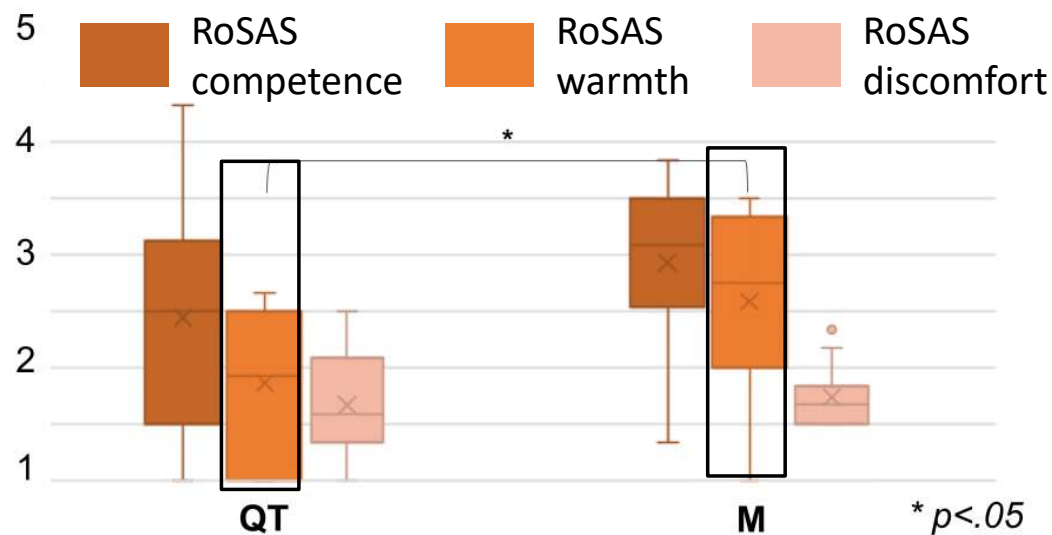
- High Openness – **O**
- High Conscientiousness – **C**
- Medium Extraversion – **E**
- Medium/High Agreeableness – **A**
- Low Neuroticism – **N**



Target Robot Personality

# Robotic Mental Well-being Coaches in the Workplace

## Results: Perception of the robots as a wellbeing coach (RQ1)



*“need to be a conversation partner first” (QT)*

*robot made them feel “a lot more engaged” in comparison to doing the exercise on their own (M)*

The Misty (M) robot was perceived more positively than the QTrobot (QT)



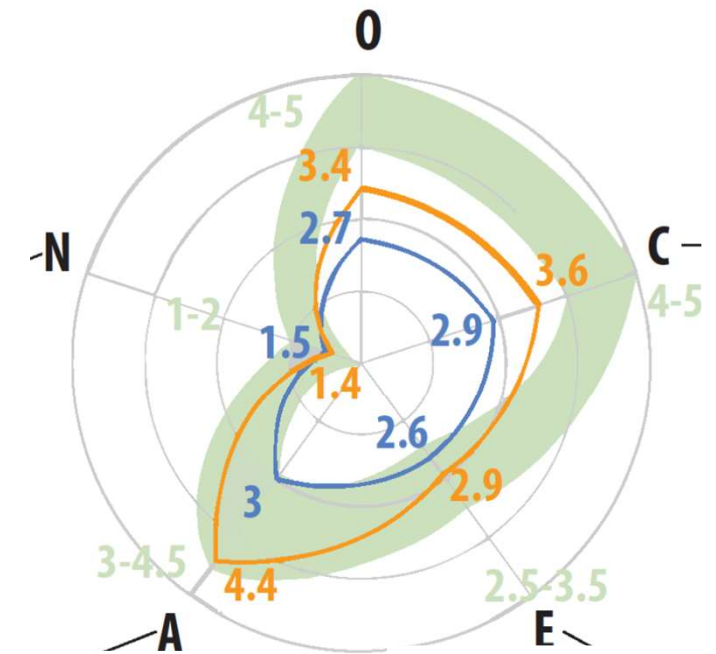
# Robotic Mental Well-being Coaches in the Workplace

## Results: Robot personality (RQ2)

*“the robot doesn’t have any personality” (QT)*

*“empathetic”, “caring”, “more introvert than extrovert”, “warm”, “calm” or “calming” and “relaxed” (M)*

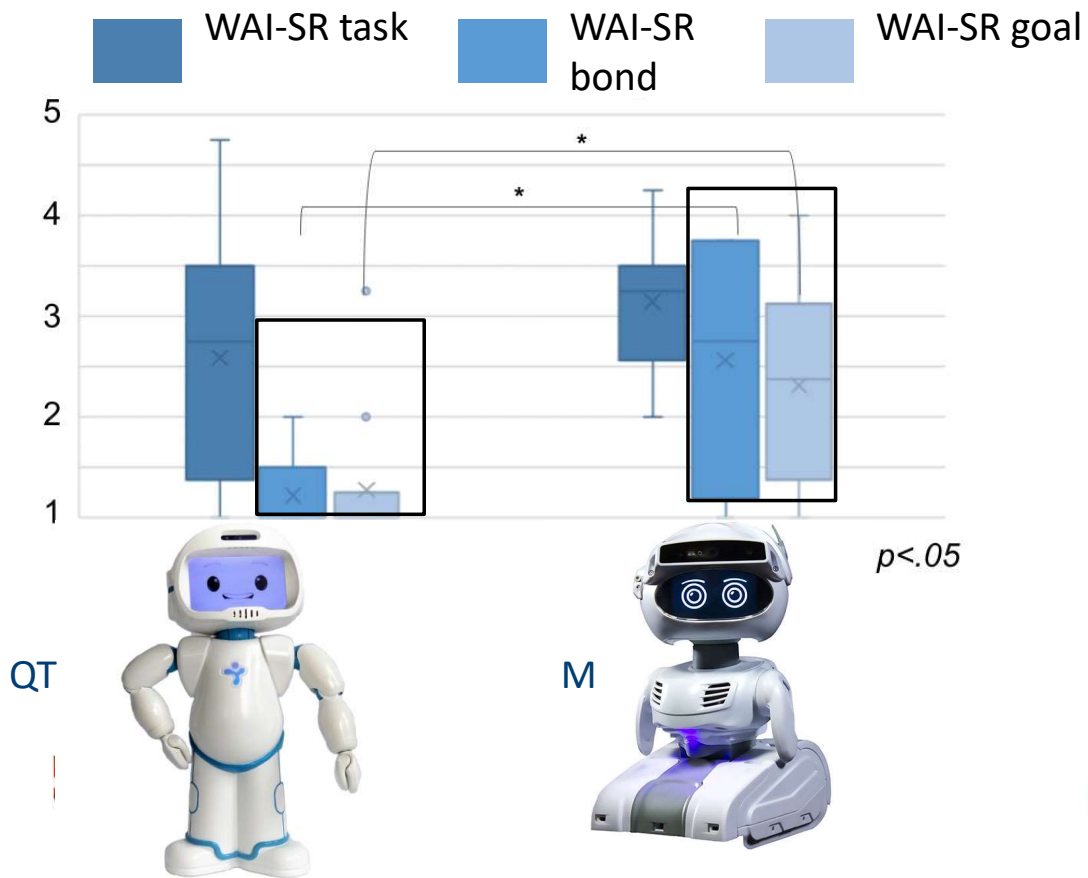
Coachees perceived **M**’s voice, gestures, and personality more positively, while coaches were more critical of the QT and many noted that it did not have a personality



Target vs. Perceived Personality of M and QT coachees

# Robotic Mental Well-being Coaches in the Workplace

## Results: Coach-coachee alliance (RQ3)



*“[the correct timing] built up my connection with the robot, and then it went and destroyed all its hard work by [talking] in the wrong places” (QT)*

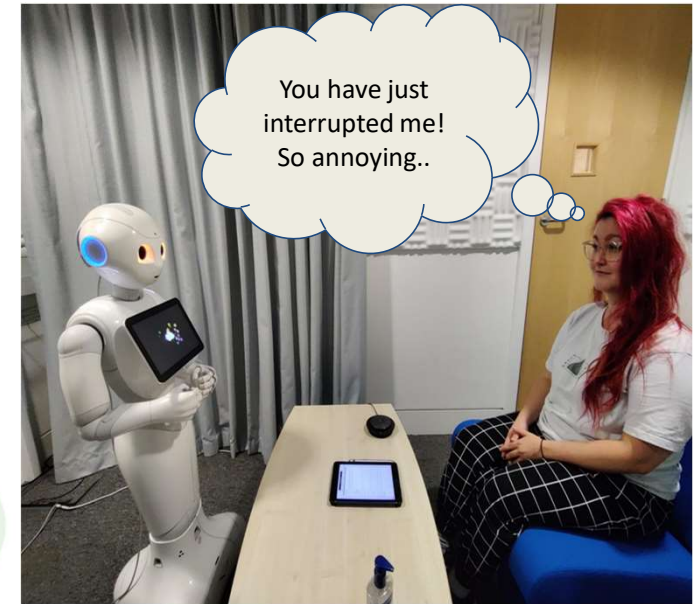
*“there’s a little emotional connection going”, “I do feel an affinity with her” (M).*

Coachees developed a stronger alliance with Misty robot (M) than with the QRobot delivering well-being exercises

## Findings & Limitations Informing the Next Study

---

- Robotic coach form / **embodiment matters**
  - Impacting perception of behaviours and personality
- The **coach-coachee alliance** is essential for successful coaching
  - but is negatively impacted by interaction ruptures





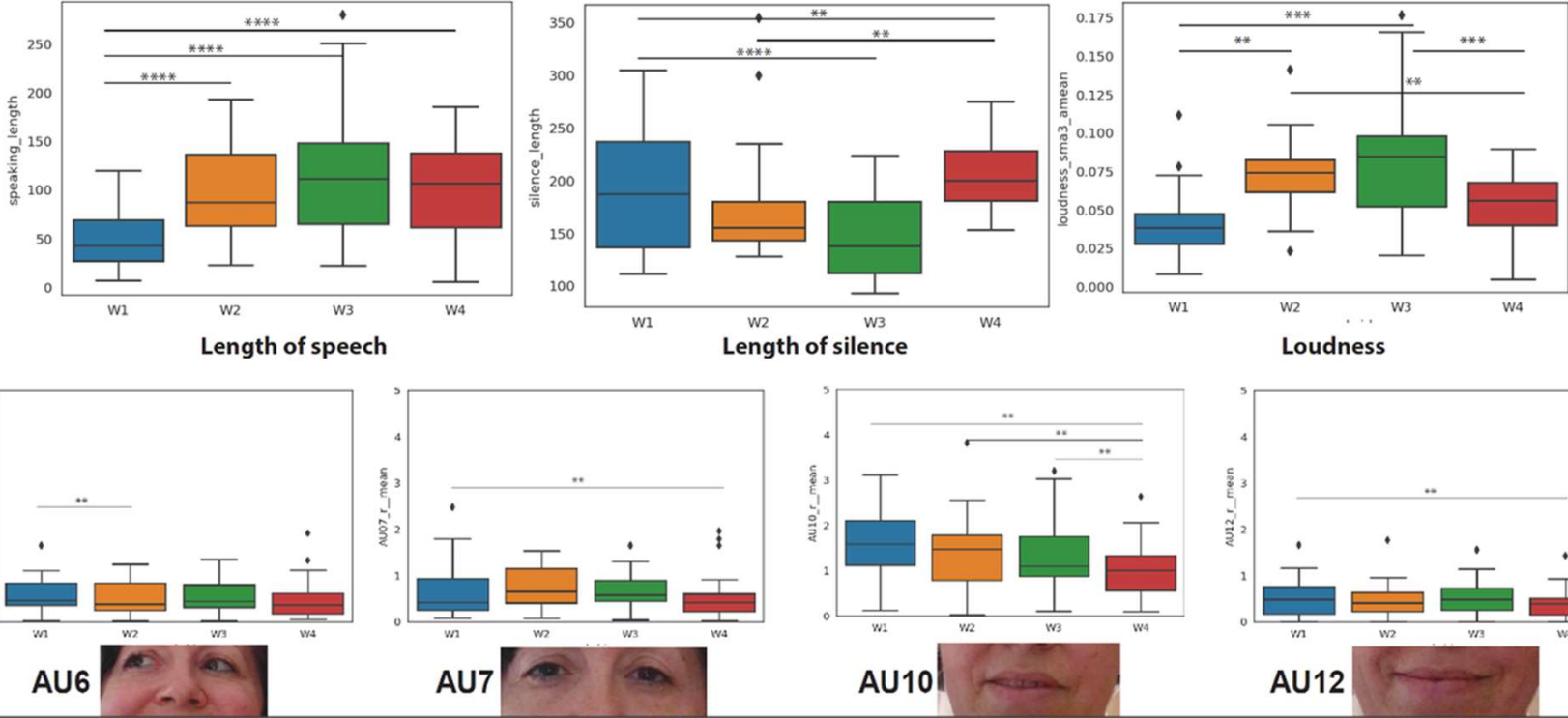
# Interaction Ruptures in Robotic Wellbeing Coaching

## ○ Data annotation

- **User awkwardness** (e.g., participants may look confused, uncertain, distressed or uncomfortable)
- **Robot mistake** (e.g., interrupting the coachee, not responding to the coachee)
- **Interaction rupture:** either the presence of user awkwardness or a robot mistake, or both



# Longitudinal Evolution of Behavioural Cues



- **Over time:**
  - **Change of** behaviour
  - **Longer speech** length
  - **Less intense** “confused” expression

Cheek Raiser

Lid Tightener

Upper Lip Raiser

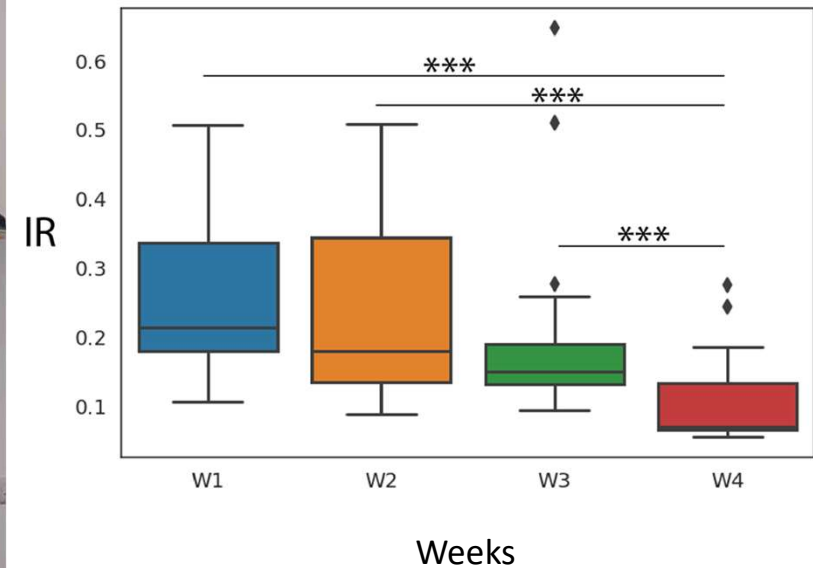
Lip Corner Puller

# Interaction Ruptures in Robotic Positive Psychology Coaching



Examples of interaction ruptures

Interaction Rupture



# Additional Limitations & Revised Objectives

---

**L1)** Negative perceptions about the robotic coach

**L2)** No longitudinal personalisation

**L3)** Occurrence of interaction ruptures

**L4)** Limited conversational capabilities

**L5)** No significant mental well-being improvement

**O1)** Interactive and responsive robotic coach

**O2)** Adaptive robotic coach over time

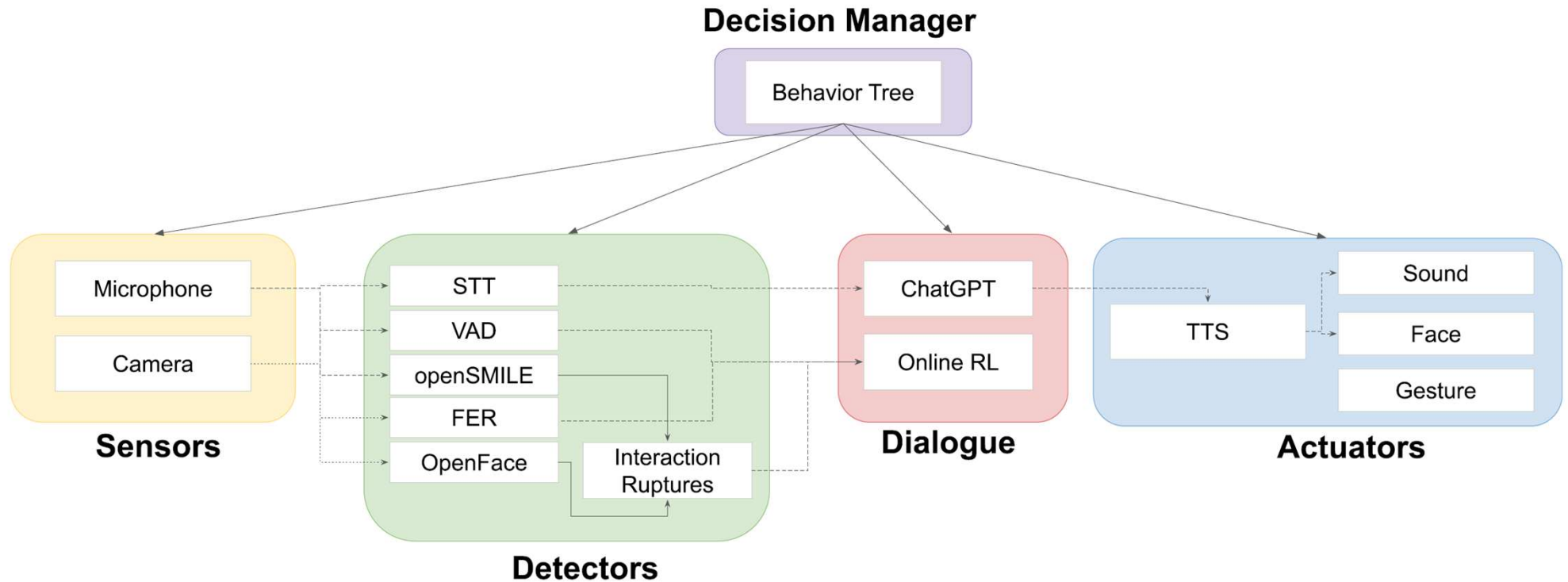
**O3)** Interaction rupture detection & repair

**O4)** LLM integration in the robotic coach

**O5)** Significant improvement of mental well-being

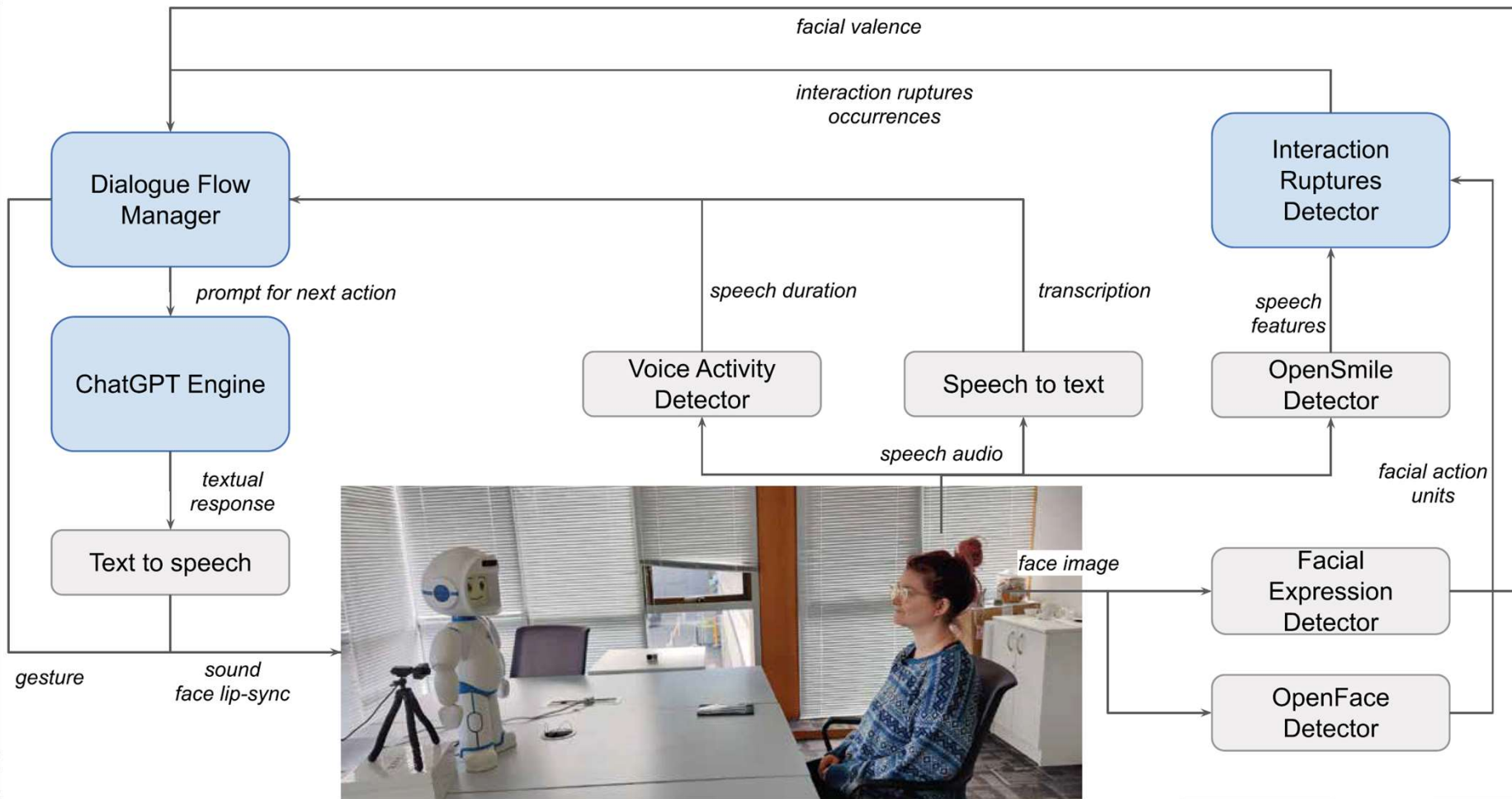
# The VITA System

Code: <https://github.com/Cambridge-AFAR/VITA-system>



VITA system includes *sensor*, *detector*, *actuator*, *dialogue* and *decision* modules using the open-source framework **HARMONI**

# Components of the VITA System



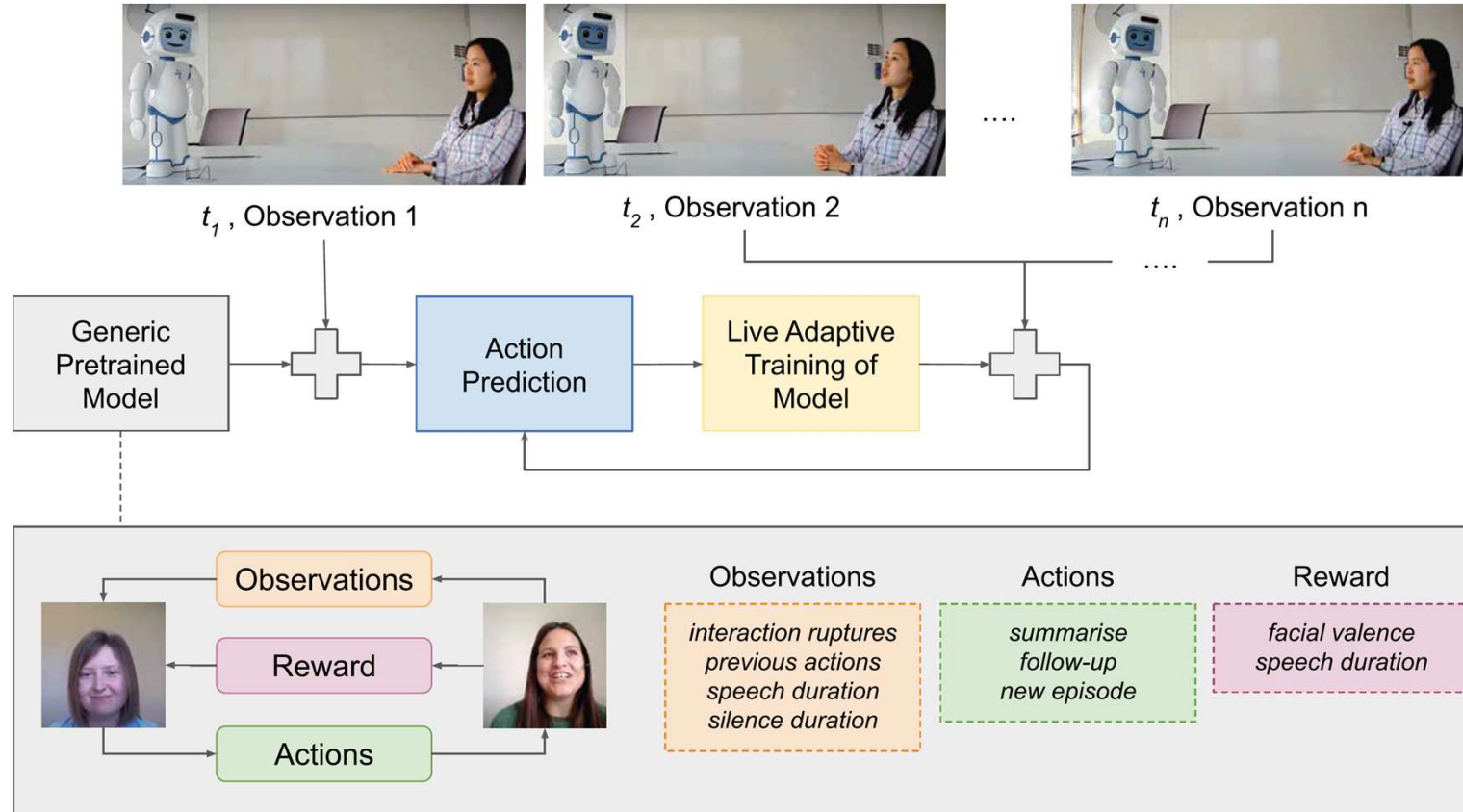
# Reinforcement Learning Pipeline

- Generic pre-trained model using a dataset we collected in our lab

- 1 well-being coach delivered 4 positive psychology exercises with 5 researchers

- Online adaptive RL:

- 3 actions
- 11-element observation space
- Reward (facial valence and speech duration)



# Evaluation of VITA



- **In the lab pilot study:** 4 researchers, robotic coach delivers 4 positive psychology exercises in 3 conditions each
  - Pre-scripted (Fig.a)
  - Generic RL (Fig.b)
  - **Adaptive RL**
- **Real-world study:** 17 employees from a tech company interacted with a robotic coach with the adaptive RL embedded over **4 weeks** (Fig.c)



# Key Findings

---

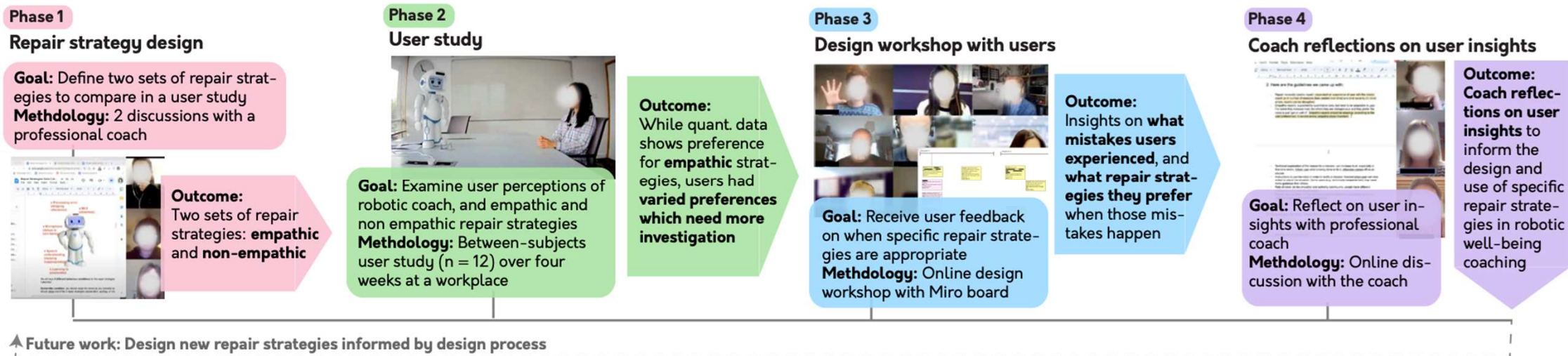
- Coachees perceived the VITA adaptive and generic configurations **more positively** than the pre-scripted one
  - **felt understood** and **heard** by the adaptive robotic coach
- VITA adaptive robotic coach **kept learning successfully** by personalising to each coachee
  - did **not** detect **interaction ruptures**
- Coachees had **significant mental well-being improvements**
- The **VITA system** is **open-source** and available on Github

# What Next? Iterative Design of Repair Strategies (sneak peek)



@ACM/IEEE HRI'24

- Problem: Closer inspection revealed **subtle interaction ruptures**
- Solution: Iterative **design of repair strategies**



## Key Takeaways (1)

---

- **Human expert** involvement and guidance matters
- **Iterative user-centred design** is crucial for user acceptance and **success**
- Use and engagement beyond the novelty effect requires **longitudinal HRI**
- Deployment in the **real world** provides **real insights** for improvement
- **Users'** perceptions of the robot and its behaviours **evolve over time**

## Key Takeaways (2)

---

- **One size does not fit all**

- Robot embodiment, personality, expressivity and speech are **inter-related** and need to be considered together for the specific **application context**
- **LLMs** are powerful for enhancing the speech processing and generation capabilities of robots, but they **need 'policing'**
- Even within the same **user** group there is **variation** → robot **learning** and **adaptation** is essential

# IEEE RO-MAN 2022 & 2023 Workshop on HRI for Wellbeing (HRI4Wellbeing)

<https://hri4wellbeing.github.io/>

HRI4Wellbeing Workshop

Objectives

Speakers

Program

Submission

Organizers

Past  
Workshop

## Human-Robot Interaction for Wellbeing Applications in the Real World

**Full-day hybrid workshop on August 28th, 2023 as part of the IEEE International Conference on Robot & Human Interactive Communication (RO-MAN 2023)**



The main topic of our workshop will be robotic applications for wellbeing in the real world, which is strongly in line with the RO-MAN 2023 theme of "Design a New Bridge for H-R-I", which seeks to address the challenges of developing intelligent robots for human health. Robots are becoming more prevalent in our society for task-oriented goals (e.g., cleaning the house, cooking a meal) and social-oriented interactions such as companionship, assistance, and coaching. We expect robots to share our daily lives in our homes, workplaces, and public spaces.

# International Journal of Social Robotics

## Special Issue: “Embodied Agents for Wellbeing”

**Call for Papers:** <https://www.springer.com/journal/12369/updates/20296154>

Visit [Nature news](#) for the latest coverage and read [Springer Nature's sta](#)



International Journal of Social Robotics

[Journal home](#) > [Journal updates](#) > Call for Papers: Embodied Agents for Wellbeing

## Call for Papers: Embodied Agents for Wellbeing



UNIVERSITY OF  
CAMBRIDGE  
DEPARTMENT OF  
COMPUTER SCIENCE  
AND TECHNOLOGY

### Guest Editors:

Micol Spitale, University of Cambridge

Katie Winkle, Uppsala University

Emilia Barakova, Eindhoven University of Technology

Hatice Gunes, University of Cambridge