Personalizing Behavior Change Intervention with Cognitive Robotics: Stage Detection & Goal Reasoning

Gloria Beraldo[1], Christian Tamantini[1], Alessandro Umbrico[1] and Andrea Orlandini[1]

[1]*National Research Council of Italy, Institute of Cognitive Sciences and Technologies (CNR-ISTC), Rome, Italy*

Abstract

Behavior change intervention (BCI) is a goal-oriented process that focuses on guiding people toward making beneficial changes in their behavior, habits, and lifestyle. This paper introduces a cognitive robotics architecture for personalizing the behavioral change process based on state detection and the evaluation of the user's attitude to choose the most appropriate BCI intervention techniques. A preliminary experimentation is presented to study the relation between the BCI qualities *self-awareness, self-efficacy* and *engagement* and the detected stage and the proposed robot's intervention.

Keywords

Behavioral Change, Personalized assistance, Cognitive Robotics

1. Introduction

Behavior change intervention (BCI) is a method designed to assist individuals in altering their behaviors to attain a positive or desired result. Digital behavior change interventions based on different persuasive computing technologies, including social robots, is a recent trend in behavior change design [1, 2]. Social robots typically have features that create empathy and attract attention, for instance, due to their embodiment shape, sounds, and visual cues (e.g., via LEDs), so they represent innovative solutions for dealing with empathy, adaptability, and constant support for personalized motivation during the whole BCI process. Through continuous and enjoyable interaction, e.g., reminders, motivational messages, and interactive activities, these robots can constitute a reliable companion towards healthier habits and improved well-being, contributing to the user's learning and reinforcement process. Finally, including social robots in BCI also enables the collection of valuable data, i.e., the insights gained during human-robot interactions can be used to refine and optimize interventions, ensuring they are aligned with individual preferences and needs.

In healthcare and assistive settings, numerous studies have already explored the use of social robots as complementary tools alongside traditional care methods. These robots have been employed to inform patients about their health status, explain the consequences of treatment, motivate patients, remind them to complete tasks (e.g., take medication, engage in cognitive and physical exercises), offer companionship, and reduce stress. Social robots have also already been investigated as persuasive technologies based on Fogg's behavior change model [3, 4], utilizing strategies like conformity, reciprocity, and authority to encourage behavioral change. However, only a few recent studies have examined the long-term effectiveness of social robots as tailored interventions for behavior change. For example, in a pilot randomized controlled trial, the social robot NAO was used to deliver a behavior change intervention aimed at reducing the consumption of high-calorie foods and drinks [5]. Similarly, a social robot was employed in another study [6] to motivate children, increasing their self-awareness and helping them develop a positive future self-image.

Workshop on Advanced AI Methods and Interfaces for Human-Centered Assistive and Rehabilitation Robotics (a Fit4MedRob event) - AIxIA 2024
gloria.beraldo@cnr.it (G. Beraldo); christian.tamantini@cnr.it (C. Tamantini); alessandro.umbrico@cnr.it (A. Umbrico); andrea.orlandini@cnr.it (A. Orlandini)
0000-0001-8937-9739 (G. Beraldo); 0000-0001-6238-2241 (C. Tamantini); 0000-0002-1184-5944 (A. Umbrico); 0000-0001-6458-5202 (A. Orlandini)

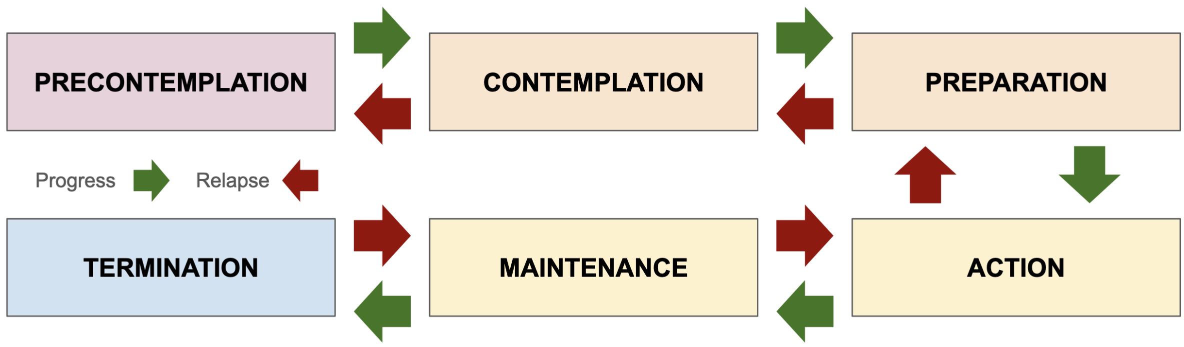

Figure 1: Stages of the Transtheoretical model

In this context, we investigate the design of a novel cognitive architecture to allow socially interacting robots to support behavioral change in individuals and provide therapists with a tool to adapt and refine the intervention strategy based on the collected information. Depending on the specific behavior the individual wishes to change (e.g., addictions, eating disorders, inconsistency in rehabilitation, or incorrect execution of rehabilitation exercises), the robot must be capable of representing knowledge about the context and the individual, perceiving and learning, possessing a Theory of Mind (ToM), planning appropriate actions, and interacting with the individual both verbally and non-verbally. Moreover, we hypothesize that social robots can tailor interactions based on individual needs and preferences to make behavior change interventions more effective based on the current BCI status [7].

The contribution of this paper is the conceptualization of a novel cognitive architecture for supporting adaptive behavioral changes via human-mediated social robots also considering ToM mediated arbitration to better adapt the proposed robotic behavior. The modularity of the architecture reflects the multi-faced abilities required by the enhanced social robots including the capability of: (i) acquiring significant data for modeling the dynamical behavior of the human over time from the patient's point of view; (ii) abstracting objective and realistic user information from the contextual analysis of the data collected with on-board and external sensors (e.g., physiological, environmental data); (iv) capturing the different ToM outputs in the context of the goal-oriented decision processes via a feasible computational model; (v) strategically reasoning about the next feedback and the kind of twofold support provided by the robot towards the patient and the caregivers/medical staff.

In particular, this paper focuses on analyzing the feasibility of personalizing the behavioral change intervention based on the BCI state detection based on the Transtheoretical model and evaluating the user's attitude to determine the most appropriate BCI intervention techniques. With this purpose, a preliminary study involving 36 participants has been conducted in the eating disorders scenario.

2. The Transtheoretical Model

The Transtheoretical Model was developed to unify various psychological approaches related to behavior change in healthcare settings [8]. This model incorporates a temporal dimension to blend processes and principles from multiple intervention theories. A particularly innovative and relevant feature of this model is its emphasis on time, conceptualizing behavior change as a progression through six distinct stages, as illustrated in Figure 1.

The stage construct introduces a temporal dimension, highlighting the dynamic nature of behavior change over time. In the **precontemplation stage**, individuals have no intention of taking action in the near future (typically defined as within the next six months). People in this stage are often either unaware or under-informed about the negative consequences of their unhealthy behaviors. They tend to exhibit low awareness of both the risks of continuing these behaviors and the potential benefits of making positive changes. Additionally, some may feel demoralized about their ability to initiate change. In the **contemplation stage**, individuals plan to make a change in the foreseeable future (usually within

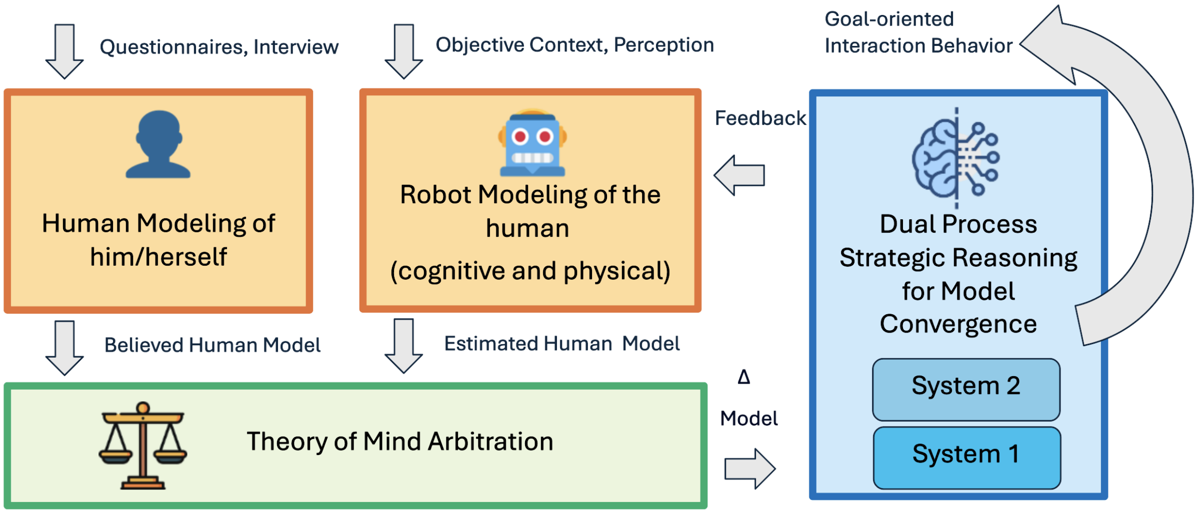

Figure 2: Architectural view of the proposed approach

the next six months) as they become aware of both the advantages and disadvantages of changing. However, they may experience a balanced weighing of pros and cons, leading to a sense of being "not ready" to take action. This indecision can cause them to remain stuck in this stage, a condition known as chronic contemplation or behavioral procrastination. In the **preparation stage**, individuals are ready to take action in the near future (typically within the next month). They recognize that changing their behavior can lead to a healthier lifestyle and are prepared to engage in action-oriented programs. In the **action stage**, individuals have recently modified their behavior and are committed to continuing with the change. At this stage, they must meet a specific criterion agreed upon by experts as adequate to lower the risk of disease or other negative outcomes. In the **maintenance stage**, individuals aim to sustain their healthy behavior by preventing relapses that could revert them to earlier stages. They should feel less tempted to revert to old habits and increasingly confident in their ability to maintain their changes. Essentially, people in this stage should develop a greater capacity to autonomously uphold healthy behaviors by enhancing their self-efficacy. This stage can last anywhere from six months to about five years. In the **termination stage**, individuals experience no temptation to revert to old behaviors and possess complete self-efficacy. They have absolute confidence in their ability to maintain their healthy habits.

3. Cognitive Architecture for Behavioral Change

The control approach presented here relies on previous works that have integrated hybrid AI technologies within a cognitive-inspired architecture for personalized and adaptive healthcare assistance and cognitive stimulation [9, 10]. We here investigate an extension introducing the Theory of Mind capabilities necessary to explicitly represent and reason about an individual's mental state. Taking into account behavior change goals, we introduce the basic constructs necessary to assess the subjective and objective states of individuals and synthesize interaction strategies suitable to promote healthy behaviors. The proposed cognitive architecture is depicted in Figure 2. It is conceived as the composition of three main processes: (a) Modeling the perspectives of human and robot agents; (b) A Theory of Mind (ToM) Arbitration; (c) The Dual Process Strategic Reasoning for Model Convergence.

Broadly speaking, the proposed architecture aims at *minimizing* the discrepancies between the two human models by synthesizing interacting/stimulation actions suitable for a particular stage of an individual within the change process. The integration of the reasoning modules in Figure 2 constitutes a cognitive loop implementing the capabilities necessary to: (i) support the decision-making of professionals within each of the stages of the Transtheoretical model; (ii) autonomously set contextual goals concerning the behavior change process, and; (iii) synthesize multi-modal interaction strategies at different time scales to influence the one's behavior in the short and long term.

The robot continuously collects individual information by integrating data from multiple sources e.g., screening procedures, physical interactions, and observations as well as physiological or environmental sensors. Similarly to a mirroring architecture [11], the collected information is used to build and maintain updated two semantically rich representations of the individual. One representation characterizes the individual perspective and collects subjective information suitable to build and infer his/her perception of him/herself and thus explicitly reason about the individual's belief about his/her state, and healthy/unhealthy behavior. One representation instead characterizes an extrinsic perspective of the individual's state and performance.

Central to the architecture is the ToM Arbitration component which is in charge of evaluating the alignment between the two models of the human to determine suitable goals for the robot and/or the professional. This module in particular analyzes available knowledge to detect possible discrepancies between the perceived, and objective/expected state of the human, infer the corresponding stage in the behavior change process, and (autonomously) determine "internal" goals to push the progress of an individual. Such goals are given to the strategic reasoning component to synthesize actions suitable to push the individual toward the next stage.

For example, an individual may declare to lead a sedentary lifestyle that he/she does not practice any sports and does not intend to take any specific action to change his/her habits. This information can be acquired by a robot through specifically designed questionnaires/interviews and used to build the subjective model of the human. The objective model could be built through analytical and verified information provided by a professional. ToM Arbitration would then compare the two models by considering a set of metrics e.g. the expected amount of consumed daily calories or daily steps, and infer the intention of the individual. The outcome of the analysis would determine a *low awareness* about the unhealthy behavior of the individual (*sedentary lifestyle*) entailing the *precontemplation stage* as the current stage of the individual's change process. According to the detected stage, the robot would (automatically) set the goal of *increasing awareness* to synthesize the socially interacting actions necessary to move the individual to the next stage.

Based on discrepancies between the two human models, possible insights of experts, as well as inferred stages and goals, the Dual Process Strategic Reasoning for Model Convergence establishes a proper goal-oriented multi-modal interaction behavior. The Dual Process Strategic Reasoning module integrates recent results in the architecture [10, 12] and relies on two sub-modules: (i) System 1 pursues opportunistic and highly reactive behavior executing safe and flexible interacting actions; (ii) System 2 pursues deliberative and goal-oriented behaviors synthesizing contextualized, personalized and robust plans.

In this regard, we design a novel AI-based robot cognitive architecture suitable to support the changing process by relying on the Transtheoretical model as background methodology. The architecture would endow social robots with the cognitive capabilities necessary to: (i) continuously monitor the state and progress of an individual; (ii) detect the current stage of the process; (iii) provide useful and contextualized insights to professionals and; (iv) synthesize suitable interaction/stimulation strategies.

It is worth noticing that we envisage the deployment of the resulting cognitive-enriched robot into a more general human-in-the-loop assistive process [13] involving professionals in each decisional choice of a "rehabilitation path". Namely, we do not aim to design a fully autonomous assistive robot. Rather, we believe that a joint combination of professional feedback and robot-based insights would represent an added value towards the design of personalized and more effective assistance.

3.1. Stage Detection and Goal Reasoning

This paper has focused on the stage detection process based on the TTT model and proposes personalized strategies to facilitate the fulfillment of the healthcare goal. With this purpose, the University of Rhode Island Change Assessment (URICA) questionnaire has been used [14] in combination with ontological reasoning to determine the stage and propose suitable intervention. Furthermore, additional questions have been added to assess the following three behavioral qualities, particularly relevant to stage detection and goal reasoning processes based on the BCI literature:

- **Self-awareness** measures to which extent a person is aware of the consequences of a particular unhealthy behavior. Within a 5-point Likert scale, values in [0, 1] mean that a person is neither worried about unhealthy behavior nor is aware of the consequences. Values in [2,3] mean that a person is somewhat aware of the unhealthy behavior and its consequences but not fully motivated to change their habits (balanced evaluation of pros and cons). Values in [4, 5] mean that a person is fully aware of unhealthy behavior and its consequences and strongly committed to behavior change.
- **Self-efficacy** measures the level of *autonomy* of a person in pursuing a healthy behavior. It is inversely proportional to the amount of stimuli a person needs to correctly achieve intermediate goals and keep the correct lifestyle. Values in [0, 1] mean that a person is not able to change his/her behavior alone and needs continuous monitoring and proactive stimuli to achieve behavioral goals/targets. Values in [2, 3] mean that a person has started achieving some results towards behavior change but is not autonomous yet. He/she would thus need stimuli to further improve his/her behavior, e.g. by increasing the difficulty of (intermediate) goals. Values in [4, 5] mean that a person is fully autonomous in achieving goals and is thus able to constantly maintain healthy behavior in autonomy, i.e. without continuous monitoring and external stimulation.
- **Engagement** measures the "direction of change" taking into account the time series of the monitored qualities of a person. Values in [0, 1] represent a *negative trend* of monitored qualities entailing a growing distance from the performance expected to achieve and/or maintain the healthy behavior. Values in [2, 3] represent a *flat trend* of monitored qualities. These values mean that a person is maintaining a substantial constant performance keeping a constant distance from the expected performance. Values in [4, 5] represent a *positive trend* entailing a decreasing distance from the performance expected to achieve and/or maintain the healthy behavior.

In [7], we propose the use of ontological reasoning that relies on an available taxonomy of BCTS [15] grouping the known interventions per TTM stage. The advantage of using such an ontology is the flexibility of extending the same framework to different domains. Indeed, the kind of interventions are expressed via high-level BCI techniques, that should be further specified through the dialogue with the experts (i.e., medical staff) and translated into robotic reactive behaviors based on the specific context.

4. Preliminary Experiments

In the present work, we conduct a further preliminary experiments to investigate the relations among the *self-awareness, self-efficacy*, and *engagement* and the behavioral change process (i.e., detected stage and proposed intervention). We hypothesize that the values of three qualities may have different meanings and relevance depending on the detected stage of the change process. Different meanings would lead to different *goals* for the strategic reasoning component and thus different interacting actions. To verify this hypothesis, we have involved 36 subjects (mean age 44.4 ± 14.6, 15M, 11F) that voluntarily accepted to participate in this study. They were required to fill out a questionnaire including the URICA tool, to determine their change stage, and specific questions aimed at quantifying their behavioral qualities, i.e. providing a subjective perception of their SA, SE, and E.

The results are shown in Figure 3. The collected data revealed that, based on the URICA questionnaire results, 18 subjects were classified as being in the precontemplation stage, 16 in the contemplation stage, and 2 in the preparation stage. From the figure, it is possible to notice the personalized intervention. A different ranking of the BCI interventions has been proposed for people that belong to the same BCI stage, that is highlighted with different size of the markers. In addition, it is worth mentioning that the same intervention can be proposed in different TTM stages but probably with different significance. For instance, consider the intervention "Comparative imaging of future outcome" has a smaller weight in the precontemplation stage with respect to other BCI strategies, while it appears more relevant in the contemplation stage.

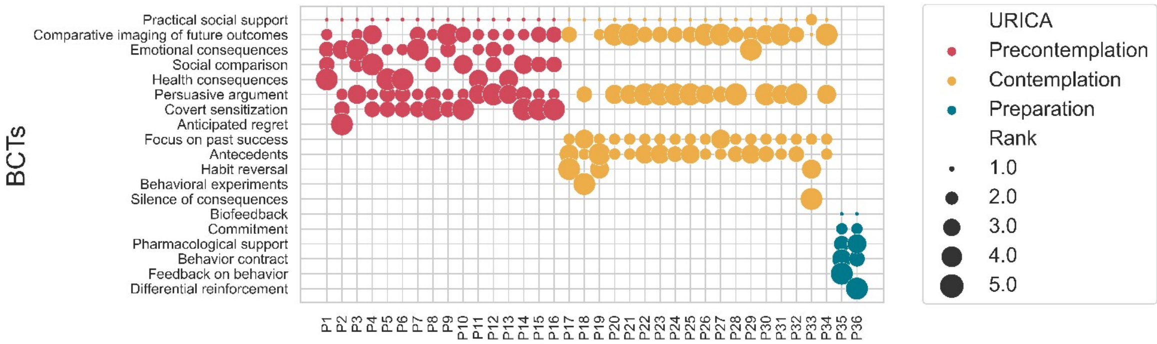

Figure 3: Scatter plots showing the distribution of the behavioral change intervention techniques across subjects and stages. The color indicates the behavioral change, the dimension of the size represents the ranking among the possible behavioral change techniques.

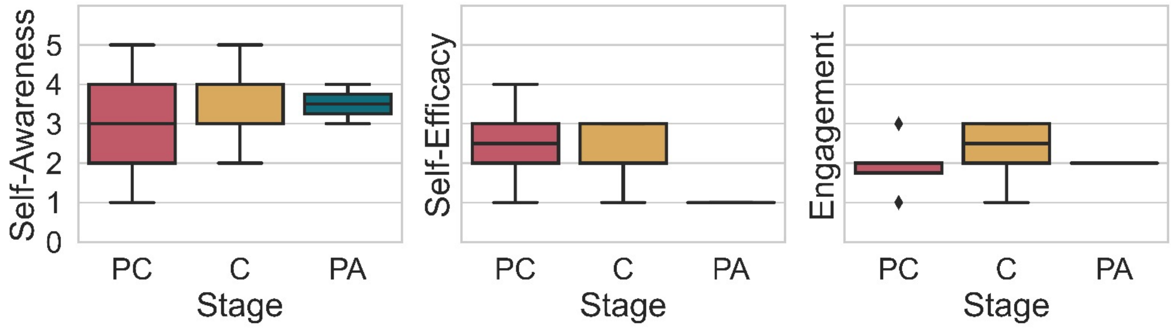

Figure 4: Boxplots showing the distribution of the *self-awareness, self-efficacy* and *engagement* over the detected behavioral stage.

As regards *self-awareness, self-efficacy* and *engagement*, the results related to the three qualities, are reported in Figure 4 per the detected TTM stage. As we hypothesize, it is possible to recognize a different trend in the 5-Likert scale answers based on the stage. The results suggest that the *self-awareness* and the *self-efficacy* are fundamental at the beginning of the process (i.e., precontemplation and contemplation stage), which correspond to higher values than in the action stage. The level of engagement gets higher in the contemplation stage. This aspect is positive given the necessity to determine progress and avoid relapses over the phases.

5. Conclusion

This paper presents a cognitive architecture for social robots designed to support behavior change interventions. Two fundamental components have been analyzed: stage detection and goal reasoning and their relations with the qualities *self-awareness, self-efficacy* and *engagement* that determine the outcome of the BCI process. Initial findings demonstrated the system's capacity to recommend interventions based on specific stages, ensuring a personalized approach for each user's needs and the current BCI stage. By prioritizing behavior change techniques (BCTs) according to emerging individual profiles, the system supports more focused and effective interventions. Future work will be focused on assessing the proposed cognitive architecture on a robotic prototype in controlled environments with the support of domain experts and real users.

Acknowledgments

This work was supported by the Italian Ministry of Research, under the complementary actions to the NRRP “Fit4MedRob” Grant (PNC0000007) and NRRP ”FAIR - Future Artificial Intelligence Research” Grant (PE0000013).

References

[1] C. Pinder, J. Vermeulen, B. R. Cowan, R. Beale, Digital behaviour change interventions to break and form habits, ACM Transactions on Computer-Human Interaction. 25(3) (2018).

[2] A. Rapp, M. Tirassa, L. Tirabeni, Rethinking technologies for behavior change: A view from the inside of human change, ACM Transactions on Computer-Human Interaction (TOCHI) 26 (2019) 1–30.

[3] B. J. Fogg, Tiny habits: the small changes that change everything, London: Virgin Books, 2019.

[4] B. Fogg, A behavior model for persuasive design, in: Proceedings of the 4th International Conference on Persuasive Technology, Association for Computing Machinery, 2009.

[5] N. L. Robinson, J. Connolly, L. Hides, D. J. Kavanagh, Social robots as treatment agents: Pilot randomized controlled trial to deliver a behavior change intervention, Internet Interventions 21 (2020) 100320.

[6] A. Triantafyllidis, A. Alexiadis, D. Elmas, G. Gerovasilis, K. Votis, D. Tzovaras, A social robot-based platform for health behavior change toward prevention of childhood obesity, Universal access in the information society 22 (2023) 1405–1415.

[7] G. Beraldo, A. Umbrico, C. Tamantini, A. Orlandini, Fostering behavior change through cognitive social robotics, in: International Conference on Social Robotics (ICSR), 2024.

[8] J. O. Prochaska, C. C. DiClemente, Toward a comprehensive model of change, in: Treating addictive behaviors: Processes of change, Springer, 1986, pp. 3–27.

[9] A. Umbrico, A. Cesta, G. Cortellessa, A. Orlandini, A holistic approach to behavior adaptation for socially assistive robots, International Journal of Social Robotics 12 (2020) 617–637.

[10] A. Umbrico, R. De Benedictis, F. Fracasso, A. Cesta, A. Orlandini, G. Cortellessa, A mind-inspired architecture for adaptive hri, International Journal of Social Robotics 15 (2023) 371–391.

[11] M. Sobhani, J. Smith, A. Pipe, A. Peer, A novel mirror neuron inspired decision-making architecture for human–robot interaction, International Journal of Social Robotics (2023). doi:10.1007/s12369-023-00988-0.

[12] R. D. Benedictis, A. Umbrico, F. Fracasso, G. Cortellessa, A. Orlandini, A. Cesta, A dichotomic approach to adaptive interaction for socially assistive robots, User Modeling and User-Adapted Interaction 33 (2023) 293–331.

[13] A. Sorrentino, L. Fiorini, G. Mancioppi, F. Cavallo, A. Umbrico, A. Cesta, A. Orlandini, Personalizing care through robotic assistance and clinical supervision, Frontiers in Robotics and AI 9 (2022).

[14] M. M. Brogan, J. O. Prochaska, J. M. Prochaska, Predicting termination and continuation status in psychotherapy using the transtheoretical model., Psychotherapy: Theory, Research, Practice, Training 36 (1999).

[15] S. Michie, M. Richardson, M. Johnston, C. Abraham, J. Francis, W. Hardeman, M. P. Eccles, J. Cane, C. E. Wood, The Behavior Change Technique Taxonomy (v1) of 93 Hierarchically Clustered Techniques: Building an International Consensus for the Reporting of Behavior Change Interventions, Annals of Behavioral Medicine 46 (2013) 81–95.