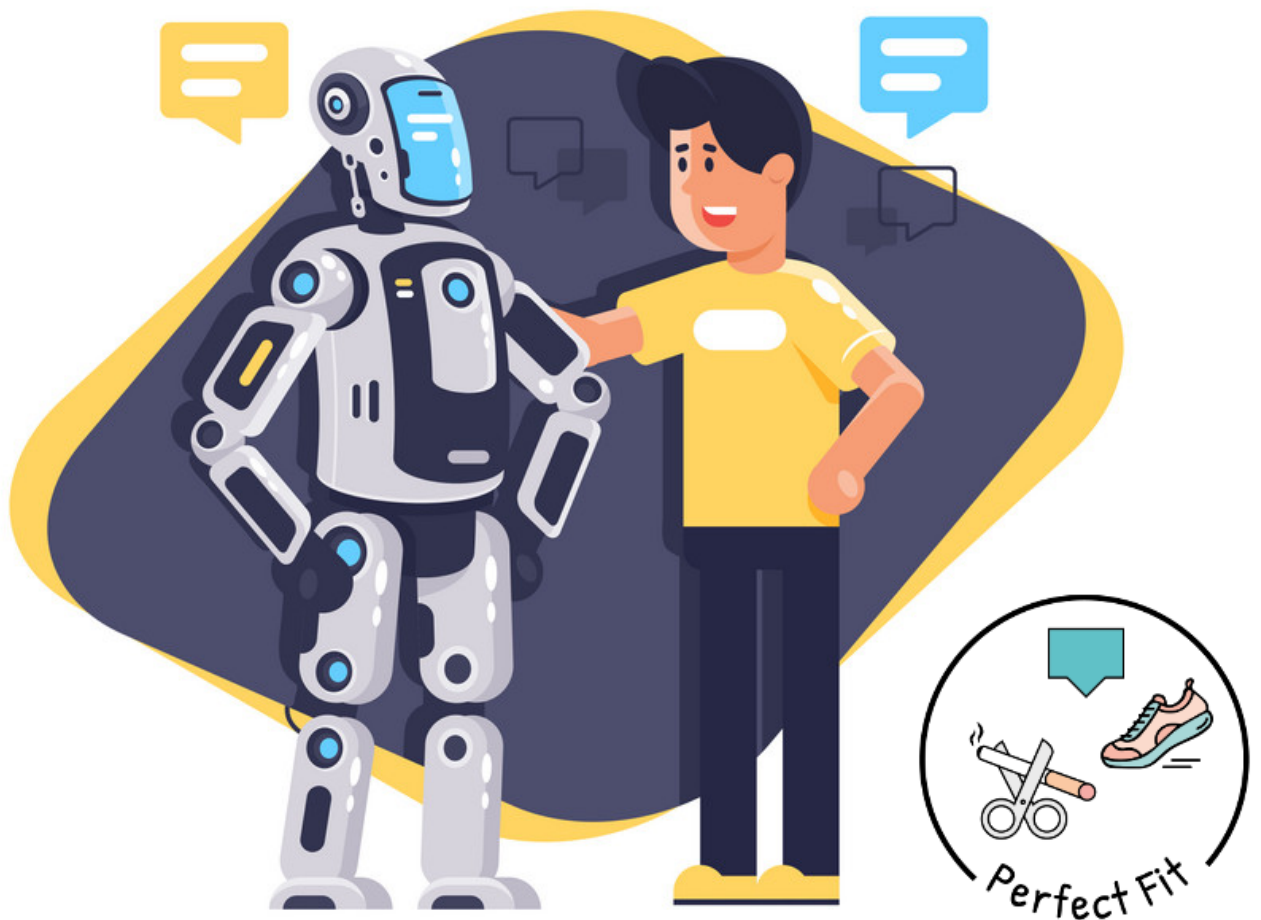


Dyadic Physical Activity Planning with a Virtual Coach: Using Reinforcement Learning to Select Persuasive Strategies

IN5000: Final Project

Andrei-Alexandru Ștefan



Dyadic Physical Activity Planning with a Virtual Coach: Using Reinforcement Learning to Select Persuasive Strategies

by

Andrei-Alexandru Ștefan

Institution: Delft University of Technology
Place: Faculty of Electrical Engineering, Mathematics and Computer Science
Project Duration: 14 November 2022 - 23 November 2023

Thesis committee: Dr. ir. Willem-Paul Brinkman - Supervisor
Ir. Nele Albers - Daily supervisor
Dr. ir. Ujwal Gadiraju - Committee member

Style: TU Delft Report Style, with modifications by Daan Zwaneveld

Contents

Acknowledgements	iii
Abstract	iv
1 Introduction	1
1.1 Motivation	1
1.2 Research question	3
1.3 Approach	3
2 Foundation	4
2.1 Goal setting theory and planning	4
2.2 Factors that influence planning	5
2.3 How to set goals and create plans	5
2.4 Virtual coaches for adaptive persuasion in behaviour change	6
2.5 Factors, values, and concerns for generating plans	7
3 Design	8
3.1 Phase 1 and details about the virtual coach	9
3.2 Plan commitment (Phase 3)	10
3.2.1 Reinforcement learning	10
3.2.2 States	11
3.2.3 Actions	12
3.2.4 Additional rules	12
3.2.5 End of the dialogue and reward function	13
3.3 Plan creation (Phase 2)	14
4 Analysis	18
4.1 Methods	19
4.1.1 Materials	19
4.1.2 Measures	20
4.1.3 Participants	21
4.1.4 Procedure	22
4.1.5 Data pre-processing and analysis	25
4.2 Results	25
4.2.1 Q1. Which aspects of the person's situation are relevant to consider for the final model?	25
4.2.2 Q2. How well can states at the end of the dialogue predict their corresponding rewards?	26
4.2.3 Q3. How well can states based on the person's situation predict states after a persuasive strategy used by the conversational agent?	27
4.2.4 Q4. What is the optimal model for persuading people to follow plans for walking?	27
4.2.5 Q5. What is the effect of (multiple) optimal persuasive strategies on persuadees' states?	28
4.2.6 Q6. How do different policies perform, compared to the optimal policy?	29
4.2.7 Exploratory analysis	31
4.3 Discussion of the results	33
4.3.1 Limitations	34
5 Conclusion and Discussion	36
5.1 Conclusions	36
5.1.1 What are factors, values, and concerns that need to be considered in a dialogue that facilitates goal setting and dyadic planning with a user?	36

5.1.2	How can a dialogue for dyadic planning with a virtual coach be designed?	36
5.1.3	How well can reinforcement learning select persuasive strategies to be used as part of a dyadic planning dialogue to make people satisfied with the dialogue, committed to the plan for taking walks, and confident in reaching their goal? . . .	37
5.2	Contributions	38
5.3	Limitations	38
5.4	Future work	38
5.5	Closing remarks	39
References		40
A Scenarios used in the expert discussions		46
B Implementation of the persuasive strategies		51
B.1	Proposing to make changes to the plan	51
B.2	Explain why planning is useful	51
B.3	Identify barriers and Deal with barriers	52
B.4	Show testimonials from other people who followed plans for physical activity	53
C Distribution of the data		55
D Detailed description of the process of selecting features and values using the G-algorithm		56
E Optimal policy		57
F Worst policy		60
G Policy when ignoring the person's situation		63
H Detailed description of the investigation into the effect of the persuasive strategies		65
I Detailed description of the thematic analysis		68

Acknowledgements

This work is part of the multidisciplinary research project Perfect Fit, which is supported by several funders organized by the Netherlands Organization for Scientific Research (NWO), program Commit2Data - Big Data & Health (project number 628.011.211). Besides NWO, the funders include the Netherlands Organisation for Health Research and Development (ZonMw), Hartstichting, the Ministry of Health, Welfare, and Sport (VWS), Health Holland, and the Netherlands eScience Center.

I would like to first thank my supervisors, Nele Albers and Willem-Paul Brinkman, for their guidance and support throughout the entire thesis. From the through feedback to the insightful discussions during the meetings, I thank you for the time and effort you have spent to bring me to this point. I wish you all the best in your future endeavours.

I would also like to thank both Ayrton and Martin for their company on campus. It felt like much less of a hassle to get things done when I knew that you were also working hard on your theses. And, of course, having someone to ask for help was invaluable. I wish Ayrton the best of luck with wrapping up his thesis and thank Martin once more for his review of my code.

I also want to thank my best friend whom I've known since high school and the friends I have met through TU Delft for the great times. I thoroughly enjoyed visiting different places in the Netherlands with you or just hanging out online and playing video games, and having something to distract me from the stress that comes with a Masters thesis is something I truly appreciated.

Finally, I would like to thank my family who supported me since I moved to the Netherlands five years ago. Be it your visits to Delft or just the video calls to chat about random things, I cherished your care and love. Thank you for giving me the opportunity to study in a different country and see more of the world.

Andrei-Alexandru Ștefan
Delft, November 2023

Abstract

Physical activity is one of the main factors that contribute to reducing the chance of chronic diseases such as cardiovascular disease, obesity, and depression, all while improving an individual's health in general. While this is the case, the fact still remains that many adults across the world do not reach the minimum recommendations for physical activity. Setting physical activity goals is one of the most common approaches found in both self-help fitness apps, or medical interventions for increasing physical activity. The issue is that goal setting on its own cannot help people become more physically active, since, if they are not committed to reaching a goal, a person will just abandon it instead. Recent literature shows that creating plans for when to perform physical activity has the potential to help people commit to reaching their goal. Being committed to following the plan is therefore important for ultimately reaching the goal, but only creating the plan is not enough, since people might abandon it. Furthermore, dyadic planning, in which a helper aids the person in creating the plan, has produced even better results than individual planning. Thus, the aim of this work is to develop a virtual coach, which plays the role of the helper in dyadic planning and motivates people to commit themselves to following the plan, so that they can reach their physical activity goals. To facilitate this process, the virtual coach, named Jamie, operates based on reinforcement learning, giving it the ability to select the best persuasive strategy to use. It does so by taking into account the person's situation (opinions of the plan and of planning in general), as well as how these opinions might change based on what persuasive strategy the agent chooses to employ. The persuasive strategies considered were: proposing to make changes to the plan, explaining why planning is useful, identifying and dealing with barriers, and showing testimonials from other people who created plans and used them to reach their goals. Through an observational study, data for the reinforcement learning model was gathered, and a model was trained on the data. Analysing the data revealed that the choice of persuasive strategy is not crucial, as all of them had similar effects on the person's situation, and these effects were small to moderate. At the same time, we saw moderate differences when comparing the situation at the beginning and end of the conversation, indicating that the combined effect of multiple persuasive strategies is needed to change a person's situation. We also investigated the effect of including or excluding the person's situation from the model. If the person's situation is disregarded, the resulting model is equivalent to following a fixed order of persuasive strategies and, through simulations, we have shown that it can change about 74% of people's situations into one where they are likely to commit to the plan. When the person's situation is included, the percentage rises to 82%, suggesting that the person's situation is also important to consider. Both of these models have advantages and disadvantages, which are discussed and addressed. Thus, this thesis provides two models for a virtual coach that can hold a persuasive dialogue in the context of dyadic planning for physical activity, which can be used as the basis for systems which target behaviour change.

1

Introduction

1.1. Motivation

Physical activity is one of the main factors that contribute to reducing the chance of chronic diseases such as cardiovascular disease, obesity, and depression, all while improving an individual's health in general [104]. Despite the recommendations given by the World Health Organization (WHO) to reduce sedentary behaviour, since "some physical activity is better than none" [20], recent studies show that, during the COVID-19 pandemic, physical activity decreased and sedentary behaviour increased [95]. According to Mattioli et al. [69], unhealthy changes in diet (such as increased alcohol consumption, or eating too much sugar) are a common way to cope with the increased levels of stress and anxiety. Considering the decrease in physical activity, the authors claim that the risk of cardiovascular disease is likely to have increased during the pandemic, especially for people who were already suffering from obesity since before the pandemic. Sedentary adults are the ones who could benefit most from interventions aimed at increasing physical activity, but they are often hard to recruit in such interventions, due to internal discouraging beliefs in their ability to exercise [55]. Thus, this work aims to provide sedentary adults with a means of becoming more physically active.

Currently, a common approach to encourage an increase in physical activity is through smartphone applications, which generally promote setting goals and self-monitoring, while providing feedback on the person's progress [71]. Setting goals is the core part of goal-setting theory, originally defined by Latham and Locke [58]. The theory mainly outlines that specific goals which are challenging for the person lead to better performance compared to urging people to "do their best", which is a generic and arbitrarily easy goal. It is possible that the goal-setting aspect of smartphone applications is a factor which contributes to an increase in physical activity for people who use smartphone apps compared to people who do not [48, 103]. However, this increase is generally only seen in the first few months, after which it diminishes, due to a decline in engagement [83].

Reduced engagement could be the effect of people abandoning the goals that they set when they downloaded the application, since goal commitment is one of the most important parts of a successful goal-setting procedure when the goals are specific and challenging [64]. According to Locke [64], there are multiple possible causes which would make people abandon their goals. First, the high complexity of a task, such as being physically active, means that people who are not experienced with doing physical activity will find it difficult to come up with strategies necessary to make meaningful progress towards the goal. Second, the idea of difficult goals puts pressure on people, in the sense that they feel the need to perform well and meet the goal no matter what, often sacrificing quality in the process, which can lead to injury (e.g. [96]). Third, the effects of negative feedback, for example a message saying that the person failed to reach a goal, are demoralising. Since not being able to achieve a goal is seen as a lack in one's own ability, it decreases people's self-efficacy (a term defined by Bandura [8], meaning people's confidence in their ability to do something, such as reaching a goal in this case). Thus, there are many factors which can cause a person to abandon their goal, so a way of diminishing their effects is necessary.

One of the strategies to increase goal commitment is developing action plans [87], which define a clear path towards the goal. By working together with the person when setting their goal and when

creating the action plan, it is possible to provide the necessary information for becoming more physically active and exercising properly, while avoiding the risk of injury. Additionally, having an action plan in place means the end goal will seem more attainable [64, 82], thus relieving the pressure put on by the goal's difficulty. Furthermore, through a technique called coping planning [86], people can think of barriers which might prevent them from following the plan. In doing so, the likelihood of negative feedback decreases, as the coping plan aims to help people come up with solutions to overcome their barriers ahead of time. Therefore, creating action and coping plans together with a person has the potential to provide them with a way of reaching their goal and prevent its abandonment.

Creating a plan together with a person can be done in two different ways. Kulis et al. [53] differentiate between collaborative planning (wherein the target person and their partner collaborate to create a plan for them to do something together, such as physical activity) and dyadic planning (which means that the target person creates a plan only for themselves, while their partner's job is to assist them, having the role of a helper). In a randomised controlled trial (RCT), Kulis et al. [53] showed that, when performing dyadic planning, both the target person and their helper managed to increase their physical activity. It is worth mentioning that the helpers in this RCT were either friends, family, romantic partners or co-workers for at least one year and met at least several times a week, so they were people close to each other, and not strangers. Similarly, Keller et al. [52] compared plan enactment in the case of individual and dyadic plans in an RCT in which the participants (both target and helper) were romantic partners. The authors found that formulating plans together with a partner leads to higher plan enactment compared to individual plans. It seems that the addition of a partner in the planning process can help anticipate barriers that the target person did not think of, or help describe the activity and what it entails more clearly [52]. Additionally, the other member of the dyad also has the role of providing social support, which motivates the target person to achieve their goal [11].

A potential problem arises when the helper in the dyad might not have the information necessary to instruct the target person on how to do physical activity properly, in order to avoid injury. Alternatively, they may not have the necessary skills to create a plan that can help the target person reach the goal. In other words, the helper might not be able to change the target person's behaviour. Fortunately, this role of helper could be fulfilled by a conversational agent. Conversational agents have been previously used to promote physical activity, an example being a paper by Bickmore, Schulman, and Sidner [12], who showed an agent that can change people's health behaviours. According to Luo et al. [67], conversational agents seem to have the potential for behaviour change with regards to physical activity. This conversation agent would then play the role of a virtual coach, guiding people through the goal-setting process and helping them create a plan for reaching their goal, while also providing information about physical activity and helping people identify potential barriers.

The virtual coach also has to take into account goal commitment. This can be seen as the virtual coach trying to persuade the person to follow the plan, since then the person is the most likely to reach their goal. However, different strategies for persuasion can be used, and they might depend on the different characteristics of people (e.g. perceived usefulness [31] of planning or confidence in being able to follow the plan [86]). For example, if a person does not perceive planning as useful for them, it might be worthwhile to explain how planning can help them. Alternatively, maybe the person is not confident that they can follow the plan, in which case the virtual coach could show them that other people managed to succeed, in order to boost the person's confidence. Factors such as confidence, and perceived usefulness can be considered a person's "state." Thus, to facilitate the interaction between a person and the virtual coach, the virtual coach needs to understand what current state the person is in, and select a way of persuading them that is the most likely to make them commit to the plan. Some consideration of what the state will look like in the future, after the virtual coach does something, should also be included. For example, thinking about barriers which might prevent them from doing physical activity might make people less confident in their ability to follow the plan. Another example is the fact that the virtual coach could help the person think of ways of overcoming their barriers only if it had already helped the person identify these barriers beforehand. An approach which takes into account states and future states to determine the best strategy to fulfill some purpose is reinforcement learning. In the context of asking people for donations, Tran, Alikhani, and Litman [99] used reinforcement learning to decide what an agent should say, based on the person's perception of donating to a charity, to determine the person to make said donation. Similarly, the virtual coach could use reinforcement learning to decide how to act, depending on the person's state, to persuade them to commit to the plan.

It is worth studying if the results regarding the potential that conversational agents have for behaviour change can be translated into a virtual coach capable of helping people develop plans, and of persuading them to follow those plans through reinforcement learning. In this study, we aimed to help sedentary people commit to the goals they set, by developing a dyadic plan that they are committed to. However, unlike previous research regarding increasing physical activity through planning interventions, in our study, creating the plan was done not with the help of a physician or specialist, but with that of a conversational agent (with the role of virtual coach). Our virtual coach helped the person set a goal, and then explore their possibilities in terms of designing a plan, rather than asking them to do this on their own, and tried to persuade people to commit to following this plan, to ensure that they would not abandon their goal.

1.2. Research question

The aim of this work is to develop a dialogue with a conversational agent, based on goal setting theory, with the purpose of setting a goal and developing a dyadic plan for reaching it. The main research question is therefore

How can reinforcement learning be used to persuade people to follow their dyadic physical activity plan in a dialogue with a virtual coach?

Sub-questions are

1. *What are factors, values, and concerns that need to be considered in a dialogue that facilitates goal setting and dyadic planning with a person?*
2. *How can a dialogue for dyadic planning with a virtual coach be designed?*
3. *How well can reinforcement learning select persuasive strategies to be used as part of a dyadic planning dialogue to make people*
 - *satisfied with the dialogue*
 - *committed to the plan for taking walks*
 - *confident in reaching their goal*

1.3. Approach

The first step towards answering the aforementioned questions was to study past research into conversational agents, goal setting theory, action planning, and reinforcement learning, in order to provide an overview of the methods used for setting goals and creating plans. Based on this and on the insights of experts in physical activity and health psychology, the first sub-question (regarding factors, values, and concerns) was answered (Chapter 2). A dialogue and a virtual coach were designed, focusing on the dyadic planning aspect of the interaction, resulting in a reinforcement learning model for the virtual coach (Chapter 3). Data for the model was gathered in an observational study, and an optimal policy for motivating people to commit to plans was created and analysed (Chapter 4). Results, limitations, and proposals for future work, as well as closing thoughts were provided (Chapter 5).

2

Foundation

The aim of this chapter is to answer the first research sub-question:

What are factors, values, and concerns that need to be considered in a dialogue that facilitates goal setting and dyadic planning with a person?

To this end, we investigated previous research regarding planning, conversational agents used in physical activity settings, goal setting theory, and reinforcement learning used to adapt to the person. In addition, we discussed with experts in health psychology and biomedical signals and systems to help further our understanding in some of these topics and make appropriate design choices. The factors, values, and concerns identified in this manner are presented at the end of the chapter as a table, and they are referenced by numbers in the following (e.g. (FVC1) for the first factor, value, or concern that was identified).

The participants in the expert discussions were a PhD student in the field of biomedical signals and systems, and a PhD student in the field of health psychology. On one hand, the first expert's background in biomedical signals and systems helped us determine the limits of physical activity per week, but also understand how to develop plans for physical activity and how to measure the intensity of the exercises. On the other hand, the second expert's input from a health psychological point of view was very important in ascertaining how people themselves create plans and what helps them commit to following these plans. The materials used during the expert discussions can be found in Appendix A. These consist of the scenarios presented to the experts.

As part of both consultations, the procedure we followed was to present the experts with scenarios that showcased the use of our system in varying contexts. For each scenario, multiple alternatives were presented and the experts were asked to pick one and argue why it is more suitable than the other ones, or argue why a different approach than the ones presented would be better.

2.1. Goal setting theory and planning

Many physical activity interventions are centered around goal setting theory. The theory was founded by Locke and Latham in 1991 [58], based on previous psychology research made by the two authors and by others. At its core, the theory states that specific and difficult goals produce better results than ambiguous goals such as "do your best" or similar formulations. Furthermore, more difficult goals lead to better performance, assuming that the person does not have conflicting goals and is committed to their goal [62].

Goal commitment is considered to be a crucial aspect of goal setting [62, 64]. Without being committed to a difficult goal, the person is most likely to just abandon it rather than keep trying to attain it, due to its difficulty. In this context, perceived goal attainability and perceived goal importance are mentioned as prerequisites of high commitment [64]. The former can be provided by planning, while the latter can be achieved by providing a convincing rationale (that could also include planning) [64]. As such, planning can be used to boost a person's commitment to their goal. (FVC1)

Besides increasing commitment to the goal, we are interested in whether or not planning can actually change a person's physical activity behaviour. Two techniques for planning, action planning [24, 30,

89, 90, 107, 109] and coping planning [24, 89, 90, 109] were identified as useful techniques for for changing physical activity behaviour. This means that following the plan allows people to change their behaviour, so we should aim to make people commit to following their plans. (FVC2)

Both coping and action plans were previously used by Scobbie, Dixon, and Wyke [86], who developed a model for setting goals and generating plans to reach them in a rehabilitation setting. The goals weren't necessarily related to physical activity, as their aim was to provide a general framework to follow in interventions that include goal setting and then action planning. This model aligns with the findings by Lenzen et al. [61], in their review regarding how the goal setting, action planning, and coping planning processes are being used in practice. Namely, there are five main steps that are followed. "Preparation" prompts the user to take an introspective look at the current situation and what they want to improve. "Formulation of goals" asks the user to formulate goals which are specific and challenging, based on the previous step. "Formulation of plan" entails developing an action plan, which describes what, when, where, and how they will work towards their goal. "Coping planning" means identifying barriers which could interfere with the action plan and thinking of solutions to overcome them, while also assessing the confidence in following the plan. The "follow-up," after a period of time, measures the progress made and gives feedback. In their research Sniehotta, Scholz, and Schwarzer [89] and Sniehotta et al. [90], also indicate that combining goal setting with action and coping planning is beneficial for aligning the person's actions with their goal. Therefore, we can conclude that developing action and coping plans after setting a goal can help guide people's attempts to reach the goal. (FVC3)

Both action and coping planning can be performed alone or with a partner. In their intervention for promoting physical activity through planning, Ziegelmann, Lippke, and Schwarzer [109] found that providing assistance to the person who is creating the plan leads to more complete plans and more physical activity in the long term (six months) than planning individually. Similarly, dyadic planning is a type of planning that involves the help of somebody else [52, 53]. Results show that it can lead to higher plan enactment [52], as well as increased moderate-to-vigorous physical activity [53]. In this context, the target (the person for whom the plan is created) and the helper (the person who assists the target) are people who are close to each other, for example friends or family [53]. Burkert et al. [21] have also shown that dyadic plans are more effective than individual plans for changing behaviour, especially in the short term. (FVC4)

2.2. Factors that influence planning

While we have established that creating dyadic plans is a good strategy to maximise goal commitment, we also have to consider personal factors that might affect how people plan or how they commit to following their plans. Such factors can be derived by looking at how people create plans for physical activity in their daily lives.

Gärling and Fujii [41] identified a person's attitude (along with other factors present in the theory of planned behaviour [3]) to have an effect on how likely a person is to make a plan (not necessarily for physical activity). They argue that overconfidence in one's abilities to perform a behaviour might negatively influence plan creation, thus causing the behaviour to be unaccomplished. Thus, ensuring that people have a positive attitude towards planning and that they are not overconfident is necessary, in order to facilitate planning. (FVC5)

While we should strive to combat overconfidence, confidence itself is also an important aspect to consider. In their research, Lorig et al. [66] asked people if they are confident that they can follow the plan, indicating that, low confidence is associated with a low likelihood of following the plan. (FVC6)

Besides attitude and confidence, Xu, Yan, and Newman [108] found that factors such as daily routines, anticipated energy levels, and past experience with physical activity influenced how people create their plans. The expert in biomedical signals and systems also suggested past experience as an important factor to consider. Therefore, routines, anticipated energy levels, and past experience should be taken into account when creating plans. (FVC7)

2.3. How to set goals and create plans

Since the target group of this work was inactive people, we assumed that they have limited experience with physical activity. Thus, it was necessary to consider how much physical activity a sedentary person can do and what suitable goals for this group are. In order to determine how to set goals, we talked to the expert in biomedical signals and systems to determine how much physical activity is reasonable for

a sedentary person. A baseline for how much physical activity people should do is the one provided by the WHO: "All adults should undertake 150–300 min of moderate-intensity, or 75–150 min of vigorous-intensity physical activity, or some equivalent combination of moderate-intensity and vigorous-intensity aerobic physical activity, per week." [20]

This amount of physical activity could be reached through walking, for example. According to Lee and Buchner [60], walking as a means of doing physical activity is appropriate for anyone, as well as easily accessible. Furthermore, walking is the most acceptable form of physical activity for people who are inactive [14], while "short bouts of activity (...) may be suitable for those initiating physical activity." [88, p. 30]. This could be accomplished through walking goals, which fall under the category of moderate-intensity physical activity when done at a moderate intensity [14]. (FVC8)

To determine how to create a plan for waking, we discussed with the health psychology expert, as part of their experience involved coaching people and creating plans. An important aspect mentioned by the expert, which is also identified by Lorig et al. [66], is that the plan should be created and owned by the individual, and not assigned to them, since this helps them commit to following the plan when they feel like they are in control. In similar fashion, Locke [64] found that when people set the goals themselves or when they participate in the goal setting process, they are more committed to the goal. (FVC9)

When creating the plan itself, planning sheets represent a tool that is sometimes used [81, 89]. Radtke et al. [81] provide an example of such a planning sheet which is designed as a table with columns that correspond to common questions often found in action plans (when, where, how, and how long). The participants are asked to fill in these columns, then formulate their plans in an "if-then" format (e.g. "If it is Tuesday, then I will go for a 30 minute walk at 6 PM"). As Hagger and Luszczynska [46] point out, the aim of these planning sheets is to help people remember and enact the behaviour they wrote down in their plan. The expert further informed us that that a visual representation of the plan is important because it provides an easy way of interpreting the plan and conceptualising an overview of what has to be done when. (FVC10)

Besides the representation, the duration of the plans is also an important aspect to consider. The expert suggested not to plan too far into the future, as this is usually difficult for people to do and detailed plans for many weeks might seem overwhelming. In line with this suggestion is the framework proposed by Gerber et al. [43], wherein the focus is on breaking a larger task into smaller parts, which can provide small visible wins throughout the process. In turn, these small wins allow people to remain committed to the end goal despite uncertainties along the way. The expert additionally mentioned that people tend to overestimate how much they can do in a certain time period. This phenomenon is often referred to as "the planning fallacy" [18, 19, 97], and it shows that people tend to predict that they will complete a task earlier than they actually finish said task [18]. Therefore, plans should be focused on the near future, and include a fixed duration for the walk. (FVC11)

2.4. Virtual coaches for adaptive persuasion in behaviour change

To go a step further than setting goals and creating plans, we can take a look at how the virtual coach can adapt the way in which it tries to persuade a person. Being able to adapt to the person means the virtual coach has a higher chance of being able to change the person's behaviour, as was shown by Maher et al. [68] and Mohan, Venkatakrisnan, and Hartzler [72]. Maher et al. [68] looked at the effects of a chatbot in a behaviour change intervention with older adults. The study was targeted at increasing physical activity and motivating dietary change through setting personalised goals for taking a number of steps and following a specific diet each week. Similarly, Mohan, Venkatakrisnan, and Hartzler [72] showed that personalised goals can motivate sedentary adults to do more physical activity. Thus, a virtual coach that can adapt to the person's needs and wants can help people change their behaviour and do more physical activity. (FVC12)

Rather than adapting the goals, we wanted to adapt the way in which the virtual coach tries to persuade people to commit to their plans. For this, we investigated the use of reinforcement learning in the context of adapting to the user. Tran, Alikhani, and Litman [99] presented an agent that acted in the context of persuading people to donate to a charity. Their agent kept track of the person's attitude towards donating, since they found attitude to be indicative of a person's willingness to make a donation. Furthermore, by leveraging several different means of persuasion, the agent learned when to try to change these opinions and when to ask if the person is willing to donate. In similar fashion,

Weber et al. [106] used reinforcement learning to adapt to the person's feedback, in the context of persuading through logical arguments. The feedback is used to infer the user's stance with regards to the logical argument provided. Their agent was tested in a study targeted at persuading people to visit a hotel, and showed that it is able to accurately predict the person's stance and adapt appropriately. Both studies used a reinforcement learning techniques to adapt to the user and successfully persuade them. Thus, reinforcement learning can be used to adaptively persuade people. (FVC13)

2.5. Factors, values, and concerns for generating plans

Based on the previous considerations for dialogues that facilitate dyadic planning with a virtual coach, we compiled the guidelines presented in Table 2.1.

Table 2.1: Factors, values, and concerns for developing a dialogue that facilitates dyadic planning

Number	Factor, value, or concern description
FVC1	planning can increase commitment to the goal
FVC2	commitment to the plan allows people to change their behaviour
FVC3	action and coping planning after setting a goal can help people by guiding them towards the goal
FVC4	dyadic planning is more effective than individual planning
FVC5	planning is more effective when people have a positive attitude and are not overconfident
FVC6	while overconfidence is undesirable, confidence in being able to follow the plan is necessary
FVC7	routines, anticipated energy levels, and past experience with physical activity are important for creating plans
FVC8	walking goals at moderate intensity are suitable for people who are sedentary
FVC9	goals and plans that are created with the person increase their commitment
FVC10	a visual representation of when, where, how, and for how long to walk is easy to understand and interpret
FVC11	planning too far into the future is difficult, since people underestimate how long their activities take
FVC12	a virtual coach that can adapt to the person can change the person's behaviour
FVC13	reinforcement learning can be used to adapt the way in which the virtual coach persuades the person

3

Design

The aim of this chapter is to answer the second research sub-question:

How can a dialogue for dyadic planning with a virtual coach be designed?

To answer the research sub-question, we took into account the factors, values, and concerns identified in the previous chapter, as well as opinions from the expert consultations. The model of the dialogue between the person and the virtual coach was designed by approaching this dialogue as a reinforcement learning problem. This would allow the virtual coach to adapt based on the situation that the person is currently in, while also taking into account how this situation might change throughout the conversation. While this chapter describes the design of the model and the interaction that was used in our study, it is worth noting that there are some differences between the study and an actual use of the virtual coach in a real-life setting, such as daily life. When applicable, we refer to how a certain aspect of the model or the interaction could be implemented in a real-life setting.

During the interaction with a person, the virtual coach has two main goals. The first goal is **plan creation**. The virtual coach chats with the person about the moments during the week when they have free time and when they feel more (or less) energetic than usual (since these were the aspects relevant for planning identified in the previous chapter), and then uses that information to create an initial plan. The second goal is **plan commitment**. In brief, the virtual coach can try to persuade people to follow the plan created previously, through several different (persuasive) strategies. These strategies are expected to change the person's situation, after which a different persuasive strategy may be used. This constitutes the main part of the dialogue, and is depicted in Figure 3.1.

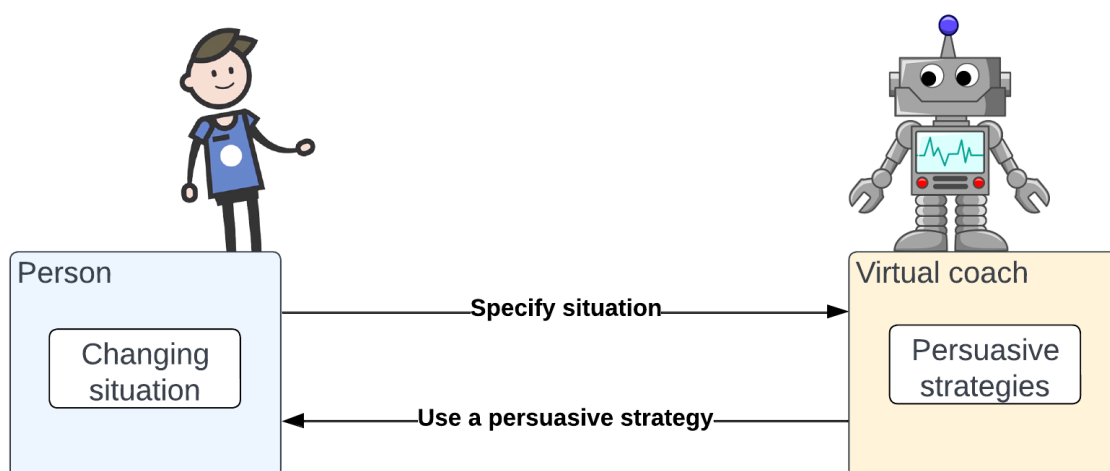


Figure 3.1: Overview of the main part of the dialogue between the person and the virtual coach

The plan commitment is handled by the underlying reinforcement learning model. For the coach to be able to select the best persuasive strategy, it is necessary to first describe what the virtual coach can see and what it can do. In reinforcement learning, these correspond to states and actions, respectively. Additionally, in order to learn what to do based on the current situation, a reward function is needed, to indicate when a certain action is better than a different one. The dialogue with the person is used to gather the data necessary for this reinforcement learning model.

Some parts of the dialogue need to follow a linear structure, to enable asking the necessary questions to begin the main part of the conversation. The overall structure of the dialogue can be seen in Figure 3.2. After a short introduction in phase 1, the virtual coach moves on to its two main goals: plan creation in phase 2 and plan commitment in phase 3.

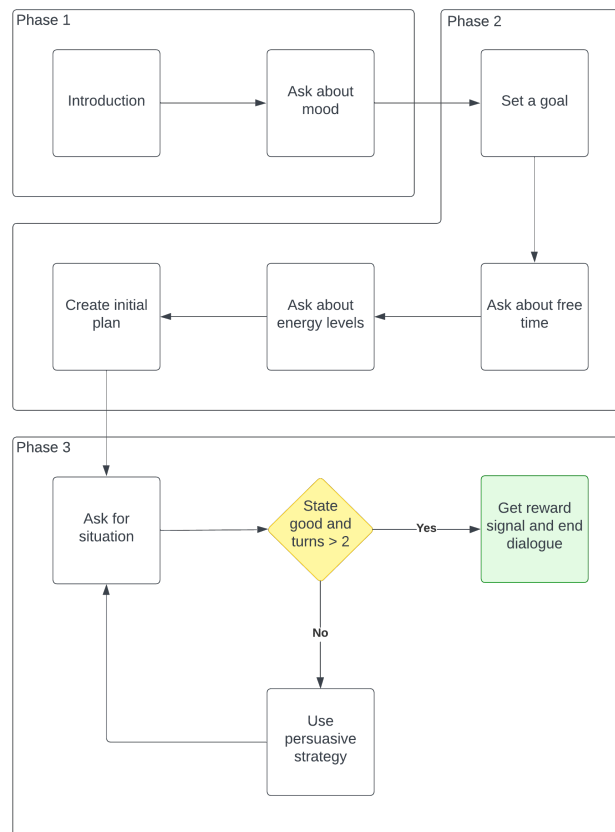


Figure 3.2: The three phases of the dialogue

3.1. Phase 1 and details about the virtual coach

The first phase of the dialogue starts with Jamie introducing itself and then asking the person about their mood. The question for the mood is based on research by Russell [84], which provides 28 options for adjectives used to describe a person's mood (Figure 3.3).

The virtual coach communicates with people through a text-based chat, as can be seen in Figure 3.3. It does not have a physical representation like an embodied conversational agent would, which means its personality is conveyed entirely through the language it uses to express itself. The virtual coach has an extraverted personality, since this personality type has a positive effect on communication satisfaction when talking about everyday activities [2]. The health psychology expert, who was asked to review the choice of language for the dialogue, additionally confirmed that the way in which the virtual coach speaks is appropriate for the context of planning for physical activity.

To convey extraversion, the virtual coach is the one who initiates the conversation, speaks in an informal manner, and sends multiple messages in a row (corresponding to a high rate of speech) [2]. Additionally, another way it conveys extraversion is through its frequent use of exclamation marks, the acknowledgement the person's responses by starting its utterances with an initial back-channel [75] (e.g. "Okay! Thanks for letting me know"). The virtual coach communicates in an informal manner, evident from the first message it sends, "Hey there! I'm Jamie, your virtual coach."

The name "Jamie" was chosen as the name of the virtual coach. As Feine et al. [34] point out, there seems to be a gender bias in the design of conversational agents, and they argue that gender-specific cues might have an effect on how people interact with the agent. Therefore, a gender-neutral name was chosen to try to diminish this effect as much as possible.

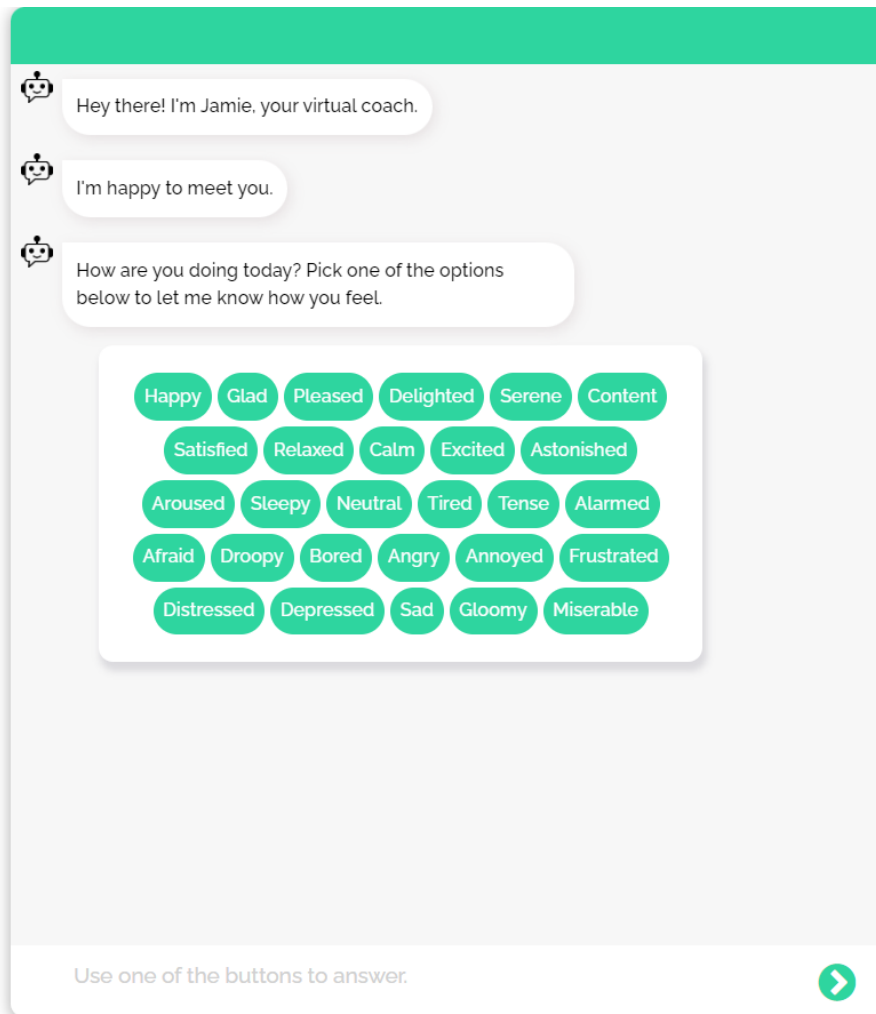


Figure 3.3: The first part of the interaction between the virtual coach and the person, including the question about mood

3.2. Plan commitment (Phase 3)

Having introduced the virtual coach, we can focus on the core part of the intervention and the main part of the conversation, plan commitment, where we present how the virtual coach employs persuasive strategies to try to make people to commit to the plan. As shown in Figure 3.1, this phase of the dialogue consists of asking the person to specify their situation, and then using a persuasive strategy.

3.2.1. Reinforcement learning

The persuasive strategies that the virtual coach can use correspond to actions in reinforcement learning. The situation that the person is in makes up a part of what is called a state in reinforcement learning.

Besides the person's situation, the virtual coach also keeps track of what persuasive strategies it has used previously, inspired by the work of Tran, Alikhani, and Litman [99]. This is the other part of the state. Therefore, the state is created by combining aspects related to the person and aspects related to the conversation. At the end of the interaction, the virtual coach asks the person to evaluate the conversation in terms of several other aspects, including their commitment to the plan. The person's evaluation of the conversation corresponds to the reward in reinforcement learning. This is a proxy measure of a person's performance. In a real-life setting, the virtual coach could ask the person which parts of the plan they managed to follow, or which ones were difficult and also adapt the plan or its difficulty accordingly.

The reinforcement learning problem of determining which persuasive strategies are best to use depending on the state can be formalised as a Markov Decision Process (MDP). An MDP is defined as a tuple $\langle S, A, R, T, \gamma \rangle$, where

- S is the state space,
- A is the action space (actions that can be done in the states of S),
- R is the reward function,
- T defines the transitions between states (the transition function),
- and γ is the discount factor, which indicates how much future rewards affect the current decision.

An agent (in this case, the virtual coach) acting in an MDP aims to learn an optimal policy π^* (a set of optimal actions to take in each possible state), which maximizes the expected cumulative sum of discounted rewards over time $E[\sum_t \gamma^t r_t]$, where t represents time steps, r_t is the reward in time step t , and γ is the discount factor. A time step corresponds to a transition between states, via an action, or, in other words, taking an action a , while in state s and moving to a state s' , while getting reward $R(s, a)$. The transition probability $p(s'|s, a)$ explains how likely it is to move to state s' by taking action a while in state s .

One possible method for learning the optimal policy is value iteration (Algorithm 1 in Hamadouche et al. [47], for example). Value iteration works by gradually building a table (called a Q-table) of state-action pairs, based on the values of a function $Q(s, a)$, that represents the quality of taking action a while in state s . During this process, the values in the table are updated at each time step t , leading to the following optimal Q-value function (called the Bellman equation [10]):

$$Q^*(s, a) = R(s, a) + \gamma \sum_{s' \in S} p(s'|s, a) \max_{a' \in A} Q^*(s', a')$$

In our study, the discount factor γ was set to 85%. The higher the discount factor is, the more value is placed on the future states. Since we wanted to account for future states, but also for current states, we chose 85% to focus slightly more on rewards in the near future, as was done by Albers, Neerinx, and Brinkman [4] for example.

3.2.2. States

The main components considered for a state are the ones related to the person's situation:

- **The person's confidence in following the plan.** The confidence in following the plan is important, as high confidence in following the plan is associated with actually completing the plan [66], and confidence can determine if people will commit to the plan or not [86].
- **The person's perceived usefulness of planning.** The perceived usefulness is another important component to take into account. In the Technology Acceptance Model (TAM) proposed by Davis [31], perceived usefulness influences the intention to use the system. So, in similar fashion, perceived usefulness of planning is expected to influence the intention to use the created plan. Creating plans that people do not perceive as useful is unlikely to make people commit to following the plan.
- **The person's attitude towards planning.** The attitude is part of the theory of reasoned action by Fishbein and Ajzen [36], and the theory of planned behaviour by Ajzen [3], wherein it influences the intention to do a behaviour. Gärling and Fujii [41] found that attitude determines how likely a person is to create a plan. As was the case for the perceived usefulness, if people do not have a positive attitude towards planning, then the plans they create are not likely to be implemented. As such, attitude towards planning is also an important aspect to look at.

Besides the person's situation, the other components of a state are the actions taken during the dialogue. Keeping track of which actions have been done is necessary, since the persuasive strategies have a limited number of times they can be used.

3.2.3. Actions

The five persuasive strategies (actions) that the agent can use are:

- **Propose to make changes to the plan.** Proposing to make changes to the plan is necessary for two reasons. First, it is important to give people control over the plan they create, since this helps increase commitment (FVC9 in Table 2.1). Second, when people make changes, they have the opportunity to take into account what they learned from the other persuasive strategies that the virtual coach used. For example, it is possible that they discovered a barrier that they did not know they had and now have to change the plan to accommodate this. Another example could be that they learned a good strategy for following plans by reading a testimonial of a different person who followed plans. Additionally, this action is based on research by Scobbie, Dixon, and Wyke [86] and by Lorig et al. [66], wherein participants were asked to create plans and then asked to redesign these plans if they did not feel confident in following them, which helped make the plans more likely to be achieved.
- **Explain why planning is useful.** Explaining why planning is useful, has the objective of trying to convince people that planning for physical activity is a meaningful first step for becoming more physically active. As the intervention is centered around creating a plan that the person is committed to, it is important to explain that this activity is useful and worthwhile.
- **Identify barriers and deal with barriers.** Identifying barriers and dealing with them are the steps of creating a coping plan [61]. The virtual coach can take these actions individually, without having to do the action of dealing with barriers immediately after identifying barriers.
- **Show testimonials from other people who followed plans for physical activity.** The last action of showing testimonials is related to Bandura's theory of self-efficacy [8], where vicarious experience is one of the methods of increasing self-efficacy (similar to confidence in our setting). Vicarious experiences can be presented by showing an example of a different person who was in a similar situation, and who managed to successfully execute the desired behaviour.

The way in which each of these actions is implemented is explained in Appendix B. Each of the actions selected were chosen so that they influence at least one of the variables that make up the user state. At the same time, some actions, such as identifying barriers, could help combat overconfidence. While the virtual coach aims to increase confidence in following the plan to make people commit to following it, overconfidence should also be avoided (FVC5), since it leads to biased decisions [85], less enjoyment from the actual performance [70], and might inhibit plan creation [41].

Actions other than proposing to make changes to the plan can all only be done once. This behaviour is enforced into the model, rather than learned. Making changes to the plan is an exception to the rule of only using persuasive strategies once, as the virtual coach can take this action twice. Having multiple chances to make changes gives people the opportunity to customise the plan to their liking, and to make use of what they learn from the persuasive strategies. On the other hand, commitment to following the plan might be negatively affected if people can modify the plan too often during the dialogue with the virtual coach. As such, this action can be taken at most two times.

3.2.4. Additional rules

Besides the rules regarding how often each of the persuasive strategies can be used, additional rules were added to enforce certain behaviours which were deemed necessary. All of these are enforced as rules, rather than learned by the reinforcement learning model.

Some actions should not be taken in certain states or before previous actions, since that would not make sense in a real-life setting. For example, it would be strange to try to deal with barriers without ever having identified relevant barriers to begin with. To prevent such situations from occurring, additional rules were added as part of the transition function:

- Rule 1: the plan cannot be modified twice in a row. In other words, an action of modifying the plan has to be followed by a different action. This rule ensures that the virtual coach does not just suggest to make modifications to the plan without trying to change the person's situation first.

- Rule 2: the action "deal with barriers" should only happen if the action "identify barriers" has already happened. This rule prevents the virtual coach from taking an action that does not make sense (dealing with barriers that have not been identified yet).
- Rule 3: the second action should be making changes to the plan if this was not done as the first action. This rule guarantees that the virtual coach does not end the dialogue without giving people at least one opportunity to make changes to the plan. It is possible that the dialogue ends after two actions, so, if making changes to the plan was not the first action, then it should be the second one. This is needed to ensure that people have the chance to customise the plan, as mentioned previously.

3.2.5. End of the dialogue and reward function

As can be seen in Phase 3 of Figure 3.2, the dialogue ends (green box) when two criteria are met (yellow diamond).

First, the state should be "good." A state is considered good when all variables which constitute the person's situation a high value. This means that a state is good when:

- the person's confidence in following the plan is high (at least an 8 on a 0 to 10 scale),
- the person's perceived usefulness of planning is high (at least an 8 on a -10 to 10 scale),
- and the person's attitude towards planning is positive (at least an 8 on a -10 to 10 scale).

Eight was chosen as the cutoff, since the scale itself would not allow for answers that are much higher. Thus, even if the person were to specify 8, 9, or 10, the specific variable would be large enough to consider it as the highest possible value of that variable.

Even if the state is good, the virtual coach should take a minimum of two actions (corresponding to having used two persuasive strategies) before it can end the dialogue. Besides the individual influence of each persuasive strategy on one or more variables that are part of the person's situation, we hypothesise that taking multiple such actions can help combat a person's overconfidence. By not setting a minimum number of actions, there is the risk of ending the dialogue right away if the state is good. Doing so would mean taking the person's high confidence at face value, thus failing to account for overconfidence.

When the state is "good" and at least two persuasive strategies have been used, the virtual coach ends the dialogue, by asking four questions, corresponding to measuring the reward function. The reward function specifies the reward gained when transitioning between states by taking an action. In our study, the reward function for states that are not at the end of the conversation is 0, and, for the state at the end of the conversation, it consists of a combination of the following:

- the person's satisfaction (*Sat*) with the dialogue (since interacting with an agent that does not satisfy the person's needs or that is annoying to talk to would lead to people deciding not to converse with the agent anymore),
- their commitment to doing physical activity in accordance with the first two weeks of the plan (C_2), as well as their commitment to doing physical activity in accordance with the entire plan (C_f),
- and confidence (*Con*) to reach the goal (since goal commitment is one of the important factors of ensuring that people don't abandon their goal [64]).

The function used in our study is $F = \frac{Sat + 0.8 \cdot C_2 + 0.2 \cdot C_f + Con}{3}$. In a real-life setting, there could be different aspects to take into account. The satisfaction with the dialogue should still be taken into account, as an unsatisfying dialogue will prevent people from conversing with the virtual coach multiple times. Instead of commitment to the plan, actual performance could be checked. Whenever the person successfully completes one day of the plan, they could chat with the virtual coach and talk about their experience, things they found difficult, or maybe even ideas for how to include some variation into the plan by changing the location, for example. The confidence in reaching the goal can be kept, but it might also be worthwhile to consider weekly adaptive goals, as they might be more suitable for motivating people to do more physical activity [1, 72], and in this case the confidence can be either about the end goal, these new weekly goals, or even both.

3.3. Plan creation (Phase 2)

Going back to the second phase, this is the point where the virtual coach gathers the information needed to create the initial plan: the person's goal, their free time, and their energy levels throughout the day. This phase starts with setting a goal. The virtual coach explains that goals should be challenging and achievable, in accordance with goal-setting theory [62]. Next, the virtual coach explains that goals are usually measured in terms of steps per day, with a common goal being to take 10000 steps each day. According to Wattanapisit and Thanamee [105], 10000 steps per day is a common recommendation and it is comparable to the WHO's recommendation of 150 minutes of physical activity at moderate intensity per week [20]. The expert in biomedical signals and systems whom we discussed with pointed out that goals higher than 10000 steps might be more suitable for more motivated individuals. As such, the person is given the option to choose 10000 steps as their goal, but they are also provided with the alternatives of 11000 or 12000 steps, in line with FVC9 in Table 2.1.

After the goal is set, the person is asked to provide the times when they are free during the week. As can be seen in Figure 3.4, people are presented with a table similar to that of a calendar. Each day is split into four different time slots. The virtual coach instructs people to click on one of the time slots for a corresponding day to indicate that they are free for at least 30 minutes during that time slot. Selected time slots turn green and display the text "Free." Additionally, people are asked to pick at least four such time slots when they have time. These time slots should span at least three different days, to ensure that people try to spread out their activity throughout the week, rather than condensing all of it into one day. Both of the experts in health psychology and biomedical signals and systems recommended that people should aim to do some physical activity every day, rather than doing a lot of physical activity on one day and no physical activity on the other days of the week.

	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
Morning (6:00 - 10:00)							
Midday (10:00 - 14:00)							
Afternoon (14:00 - 18:00)							
Evening (18:00 - 22:00)	Free	Free	Free	Free			

I'm done selecting time slots when I am free.

Figure 3.4: The table for specifying free times

Besides the times when people are free, the initial plan also considers the times when people have more energy than usual (FVC7). Rather than asking for a person's energy for each day of the week, the virtual coach only distinguishes between weekdays and weekends, since it might be difficult for people to recall their energy levels at every time and in every day of the week. This can be seen in Figure 3.5. Similarly to the table for free times, each day is split into four different time slots. People are asked to specify how much energy they have, compared to their average energy. They are not asked to provide this for every week of the plan, but only to indicate how their energy levels typically shift during a regular weekday and during a weekend day. When clicked, each slot displays the corresponding level and turns a different colour, to help easily distinguish the different options. The colour palette was created using an accessible colour palette generator [102], and checked using a colour palette checker which displays how the most common colour blindness types would see the different colours [77].

Weekdays	Morning (6:00 - 10:00)	Midday (10:00 - 14:00)	Afternoon (14:00 - 18:00)	Evening (18:00 - 22:00)
None at all	None at all			
Less than usual		Less than usual		
An average amount			An average amount	
More than average				More than average
Much more than average				
Weekends	Morning (6:00 - 10:00)	Midday (10:00 - 14:00)	Afternoon (14:00 - 18:00)	Evening (18:00 - 22:00)
None at all				
Less than usual	Less than usual			
An average amount		An average amount		
More than average			More than average	
Much more than average				Much more than average

I'm done specifying my energy levels.

Figure 3.5: The table for specifying energy levels

When the goal, the free times, and the energy levels are all specified, the initial plan is created and displayed on the screen. The plan is presented as a visual weekly planner, similar to a calendar, as can be seen in Figure 3.6. We chose this means of depiction, since the content and time of the plan need to be represented in a manner that makes the plan easy to understand (FVC10). There are several alternatives for this, such as text, images, or both, but the easiest to understand is an audio-visual representation [59], from which we chose to present only the visual, since including audio was not suitable for our choice of a chat-based medium.

In Figure 3.6, the plan shown spans three months. In fact, every plan created will span three months. The rationale behind setting a fixed duration for the plan is given by the health psychology expert, who explained that interventions typically do not aim to coach people for a very long time, as the goal is to enable people to plan for themselves eventually. Furthermore, the expert informed us that plans typically last two to three months for goals that are set half a year later. From FVC11, we know that planning too far into the future is difficult, so setting a restriction on how far plan can stretch is necessary. In terms of the goal, it was also found that distal goals make it easy to postpone effort [72], so we place the goal far enough for people to have time to reach it, but not so far as to postpone effort. Thus, the goal is set at six months into the future.

Figure 3.6 also shows the first two weeks of the plan described in detail: for each day of the week, it is indicated when the walking activity should be done, as well as how long it should be done for. The two weeks after those display how much and how often the person should walk in that specific week. Then, the second and third months indicate how much the person will have to walk in the final week of that month. The choice of representing only the first two weeks in detail is motivated by the fact that showing more than a few weeks might be overwhelming for people (FVC11). For weeks after the first two, the decision was to only display the total duration for that week, and the number of walks in the week, to give an abstract idea of what the plan will look like and what people would have to be able to do in the end.

Week 1	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
Morning (6:00 - 10:00)							
Midday (10:00 - 14:00)							
Afternoon (14:00 - 18:00)							
Evening (18:00 - 22:00)	Walk 30 minutes			Walk 30 minutes		Walk 30 minutes	Walk 30 minutes
Week 2	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
Morning (6:00 - 10:00)							
Midday (10:00 - 14:00)							
Afternoon (14:00 - 18:00)							
Evening (18:00 - 22:00)	Walk 35 minutes			Walk 35 minutes		Walk 35 minutes	Walk 35 minutes
Week 3	Walking for 3.0 hours, distributed across 4 time slots						
Week 4	Walking for 3.0 hours, distributed across 4 time slots						
Month 2	Walking for up to 5.0 hours per week, distributed across 5 time slots each week						
Month 3	Walking for up to 6.5 hours per week, distributed across 6 time slots each week						

Figure 3.6: Depiction of a plan

As was already noted, plans contain multiple walking activities, each composed of two important characteristics: the **duration** of the activity and the **scheduled time** of the activity.

Duration

It is important to first note the representation of the walking activities within the plan. Thus far, we have discussed the goal in terms of steps per day, while the plan shows a duration in minutes, rather than steps, to give people an indication of how long an activity would take (FVC11). We consider 100 minutes of walking to be equivalent to taking 10000 steps, since walking at a rate of 100 steps per minute corresponds to moderate intensity [101] (FVC8). Additionally, the durations presented in the plan are rounded: durations in minutes are rounded to the nearest multiple of five (e.g. 32 minutes becomes 30 minutes), while durations in hours are rounded to the nearest half hour (e.g. 2.7 hours becomes 2.5 hours). Research has shown that rounding numbers helps people recall the number with better accuracy [76], which aligns with FVC10 in Table 2.1.

The duration depends entirely on the goal that the person set at the beginning of the dialogue. We make the assumption that every person (including inactive people) takes a minimum 2000 steps per day [100]. Then, the goal of 10000 steps per day corresponds to an increase of 8000 steps. In minutes, that is 80 minutes per day, in every day of the week, meaning 560 minutes per week. By starting with a baseline of 120 minutes of walking activities in the first week and increasing by 20 minutes each week, the person would reach their goal of 560 minutes in the 22nd week. Recall that the goal is set six months into the future, meaning 24 weeks. Thus, a linear increase of 20 minutes per week can reach this goal. A similar argument can be made for the goals of 11000 and 12000 steps, with the only difference being that the increase per week should be 22.5 or 25 minutes, respectively.

Scheduled time

Closely related to the duration is the scheduled time of the activity, representing moment when it should be done. When determining these moments, we take into account the time slots when people are free

and their energy levels, both of which were specified in the second phase of the dialogue.

When indicating the time slots when they are free, people are asked to specify at least four time slots when they are available for at least 30 minutes. Considering that the first week of the plan consists of 120 minutes of walking, a minimum of 30 minutes are required if four time slots are used. If more than four time slots are specified, energy levels can also be taken into account. This can be done as follows:

- from the list of time slots when people are free, first select all the time slots when the person specified their energy level as "Much more than average",
- then, if four time slots have not yet been picked, repeat this for each lower energy level, until four time slots are selected,
- if, at any point, selecting all selecting all the time slots of an energy level would result in having more than four time slots selected in total, randomly pick the number of time slots needed to reach four in total from the corresponding energy level.

Note that it is not necessary to consider energy levels when only four time slots are specified by the person. In that case, those four time slots need to be selected regardless of energy levels.

The last consideration in terms of the scheduled time of the activity is how specific this time should be. The plans shown to the health psychology expert had an entire time slot occupied by the activity, rather than a specific hour within that time slot. For example, if a walk was scheduled for the evening, the entire evening (described arbitrarily as 18:00-22:00) was reserved for this purpose. The expert agreed with this representation, specifying that it is clear (FVC10) and suitable to make sure people do not schedule too many activities (which can include activities other than walking) in the same time slot (FVC11). This can also be seen in Figure 3.6, where the day is split into four blocks, each spanning a duration of four hours.

4

Analysis

This chapter answers the following research question:

How well can reinforcement learning select persuasive strategies to be used as part of a dyadic planning dialogue to make people

- satisfied with the dialogue,
- committed to the plan for taking walks,
- confident in reaching their goal.

A reinforcement learning model for determining which persuasive strategy to pick based on the state was trained. Its purpose was to optimise the three factors mentioned above - satisfaction, commitment, and confidence - corresponding to the reward. To analyse the performance of this model, several different aspects were investigated through several sub-questions.

The state considered in the reinforcement learning model has components related to the person (confidence, perceived usefulness, and attitude), and components related to the conversation. Considering the components related to the person, it is possible that some of these features are more important to consider than others. For example, since both perceived usefulness and attitude refer to planning, it might be the case that one of them captures a person's situation regarding planning better than the other. The state space of the reinforcement learning model is quite large, therefore, by selecting only a subset of the person-related features, it would be possible to reduce the size of the state space. A smaller state space is favourable in reinforcement learning, since having less states means a lower chance of there being states for which the optimal policy is to pick an action uniformly at random due to there not being enough data to determine an optimal action. Additionally, removing some of the features would mean not needing to ask people many repetitive questions during the dialogue, which would prevent breaking the flow of the conversation.

Furthermore, it is necessary to determine how many values each feature should have. For example, maybe a person's confidence can be fully captured using two values (such as low and high), but their attitude requires three values (such as negative, neutral, and positive). Thus, the first research sub-question is:

Q1. Which aspects of the person's situation are relevant to consider for the final model?

After determining what aspects of the person's situation to include in a state, it was possible to investigate if the state provides a good indication of the reward. A reward is associated with the final state of a person's dialogue. If this state could accurately predict what the reward is going to be, it would mean that the state gives a good indication of whether or not the person is committed to the plan, confident that they can reach their goal, and satisfied with the dialogue. The second research sub-question is:

Q2. How well can states at the end of the dialogue predict the reward?

While states at the end of the dialogue might provide an insight into the reward, states throughout the dialogue could also give insight into future states. More specifically, a state could predict what state the person will be in after the virtual coach uses one of its persuasive strategies. Being able to predict the state after a persuasive strategy is favourable, since it would then be possible to pick a strategy that has a beneficial impact on the person's future state (and eventually, the reward), for example, by improving their attitude towards planning. As such, the third research sub-question is:

Q3. How well can states based on the person's situation predict states after a persuasive strategy used by the conversational agent?

All of the questions discussed thus far give an insight into different aspects of the optimal model for selecting persuasive strategies based on a person's state. Besides the considerations regarding features, states, and future states, the optimal model also includes the optimal policy, which indicates which persuasive strategy should be used, and whether this persuasive strategy depends on the state that a person is in. The fourth research sub-question is:

Q4. What is the optimal policy for persuading people to follow plans for walking?

Once the optimal policy is determined, we can further investigate it. By simulating how a person's state changes when following an optimal policy, it is possible to see how many persuasive strategies are necessary to reach the end of the dialogue. Additionally, it is possible to check what percentage of people would reach states in which they are likely to be persuaded to commit to following plans for physical activity, as this would indicate a successful persuasive dialogue. The fifth research sub-question is:

Q5. What is the effect of (multiple) optimal persuasive strategies on persuadee's state?

An optimal policy is computed by selecting the action which has the highest Q-value in the Q-table. It indicates which action to take in a specific state. Therefore, computing a policy in a different way would potentially mean that a different action should be done in a certain state. The worst policy, for example, picks the lowest Q-value in the Q-table. By comparing these two policies, it is possible to gain insight into the effects of the actions chosen on the reward at the end of the dialogue. If both policies have a similar reward at the end of the dialogue, it means that the actions have similar effects. The sixth research sub-question is:

Q6. How do different policies perform, compared to the optimal policy?

4.1. Methods

To answer all the research questions, an observational study was run in June and July 2023, to gather data for building the model. Before starting the study, all the details concerning the setup and the analysis were registered with the Open Science Framework (OSF) [93]. The TU Delft Human Research Ethics Committee (HREC) approved the setup of the study (HREC reference number: 3089).

4.1.1. Materials

The platform Prolific¹ was used for recruiting participants, whose data was collected through two different means. For the pre-screening and post questionnaires, Qualtrics² was used, both to create the questionnaires and to gather the responses. For the conversational session, a MySQL³ database was used to store the participant's responses which were part of the primary or secondary measures, while a PostgreSQL⁴ database was used to track the entire conversation. Both of the databases, and the virtual coach developed with Rasa⁵ version 3.2.8 were hosted on Google Compute Engine⁶. The data gathered is available at [92], and the code for the virtual coach is available at [91].

Additionally, data from previous research was used for one of the persuasive strategies. The dataset from Albers et al. [5] was used, as it also contained testimonials from people who found planning to

¹<https://www.prolific.com/>

²<https://www.qualtrics.com/>

³<https://www.mysql.com/>

⁴<https://www.postgresql.org/>

⁵<https://rasa.com/>

⁶<https://cloud.google.com/compute>

be useful. These testimonials were used to formulate the messages used in the action of showing testimonials of other people who used plans to reach their goals.

4.1.2. Measures

Primary measures

The primary measures are the ones used to directly answer the research questions. The following were measured:

1. Variables which are part of the state (confidence in following plans, perceived usefulness of planning, and attitude towards planning)
 - (a) **Confidence in following plans.** This variable was measured by adapting the question used to measure confidence in Morony et al. [73]. The question was reformulated to reflect confidence in following the plan, by replacing the subject of the original question with the confidence in following the plan. Values ranged from 0 to 10.
 - (b) **Perceived usefulness of planning.** This variable was measured by asking the question "To what extent do you think planning can help or hinder you in terms of taking walks?" on a -10 to 10 scale from "I think it can hinder me very much" to "I think it can help me very much," based on the technology acceptance model (TAM) by Davis [31]. As we wanted to use a single question for each measure, the factor loadings for the questions used to measure aspects of the TAM calculated by Larmuseau, Desmet, and Depaeppe [57] was used to determine which question of the TAM to pick.
 - (c) **Attitude towards planning.** This variable was measured by asking "Making plans for taking walks is..." on a -10 to 10 scale from "Bad" to "Good," based on the guidelines presented by Francis et al. [40].
2. Variables which are part of the reward (satisfaction with the dialogue, commitment to the first two weeks of the plan, commitment to the whole plan, and confidence in reaching the goal)
 - (a) **Satisfaction with the dialogue.** This variable was measured by asking "How satisfied or dissatisfied were you with the dialogue?" on a -10 to 10 scale from "I am very dissatisfied" to "I am very satisfied." The question was based on research by Hiraoka et al. [49], wherein the authors ask for subjective satisfaction, and adapted to fit a scale with diverging endpoints.
 - (b) **Commitment to following the plan.** The question asked to measure commitment to following the first two weeks of the plan was adapted based on research by Butryn et al. [22]. The question was adapted to the domain of the study, by specifying that the behaviour is taking walks by following the plan for the first two weeks. The same question, but worded to reflect the entire duration of the plan, was asked to measure commitment to the entire plan. The scale ranged from -10 to 10 ("I strongly disagree" to "I strongly agree").
 - (c) **Confidence in reaching the goal.** This variable was measured by adapting the question used to measure confidence in Morony et al. [73]. The question was reformulated to reflect confidence in reaching the goal, by replacing the subject of the original question with the confidence in reaching the goal. The scale ranged from 0 to 10.

Demographic measures

Demographic measures are the ones which are gathered for participant selection:

1. **Measures from Prolific.** People's age, gender, and weekly physical exercise were gathered from Prolific, along with their number of completed studies and approval rate.
2. **Leisure time physical activity.** The amount of physical activity people do was measured by asking them to fill in the Godin-Shephard leisure-time physical activity questionnaire [45].
3. **State of change for physical activity.** This was measured using an adaptation based on the short form for exercise by Nigg [78]. The original question was adjusted to provide the definition of what physical activity is according to the WHO [20] and then asking people if they are physically active according to this definition.

Secondary measures were the ones gathered for the exploratory analysis:

1. **Frequency of planning activities.** Based on an article on measuring behavioural frequency published by Qualtrics [42], two questions about frequency of plan creation were formulated: "In the past 3 months, how many plans for things you need to do in the future did you create?" and "In the past 3 months, how many plans for taking walks did you create?"
2. **Enjoyment of planning activities.** Two questions regarding enjoyment of creating plans are asked, adapted based on The Groningen Enjoyment Questionnaire by Stevens et al. [94] and reformulated to reflect enjoyment of creating plans rather than enjoyment of doing physical activity.
3. **Opinions regarding the interaction and the agent.** The interaction between the person and the virtual coach was evaluated in terms of the person's opinions regarding the coach, such as believability, usability, performance, and opinions regarding the interaction, such as user acceptance, user engagement, and user-agent interplay. The questions asked were the ones in the short form of the Artificial Social Agent (ASA) Questionnaire [39].
4. **Motivating and demotivating aspects of the interaction.** Participants were asked to answer the following questions: "What part of your conversation with Jamie was the most motivating for you, in terms of following the plan?" and "What part of your conversation with Jamie was the most demotivating for you, in terms of following the plan?"

Other measures which were gathered but not used in the analysis are detailed in the OSF form [93]. These include people's takeaways from the testimonials they were shown, their opinions of how planning can help them become more physically active, and descriptions of their action and coping plans.

4.1.3. Participants

Sampling procedure

Participants were recruited through Prolific⁷ in June and July 2023 and paid based on the payment rules on Prolific (minimum £6 per hour). People who were registered in a similar study [32] at the time this study was run were excluded from participating. Based on the answers to the pre-screening questionnaire, participants were invited to the conversational session. The target group of this study was people who are inactive or unfit. To determine if a person is inactive or unfit, we considered their weekly amount of exercise (entered on Prolific), their state of change for becoming more physically active, and their amount of exercise according to the Godin-Shephard leisure-time physical activity questionnaire. Thus, only people who answered "Never (0-60 minutes per week)" to the question "How often do you engage in physical exercise per week?" on Prolific and were in the contemplating or preparing state of change for physical activity were eligible for participation. Of these participants, the ones which qualified as "Insufficiently Active/Sedentary" in the Godin-Shephard leisure-time physical activity questionnaire were recruited first.

Before running the study, we decided to add people who qualified as "Moderately Active" in the Godin-Shephard leisure-time physical activity questionnaire if there were not enough people who qualified as "Insufficiently Active/Sedentary." During the experiment, it was indeed the case that there were not enough sedentary people, thus moderately active people were also recruited.

Based on the list of eligible participants, people who were invited to participate were selected to produce a balanced sample across gender. Additionally, we also wanted to balance the participants based on their current situation (their confidence, perceived usefulness, and attitude), in order to try to make all possible starting states of the reinforcement learning model have an equal number of samples.

Considering the possibilities for the person's situation part of the state, there are 12 possible starting states, corresponding to all possible combinations of three levels for confidence, two levels for perceived usefulness, and two levels for attitude which were considered initially. It was not possible to obtain a perfectly balanced distribution across starting states and there was no time left to invite new participants (since the duration of the study was restricted to one week), thus participants who had completed the pre-screening questionnaire and were eligible were invited to participate in the conversation session, regardless of starting state, but balanced according to their gender. The study included 114 participants, and their characteristics can be seen in Table 4.1.

⁷<https://www.prolific.co/>

Sample size and power

Guidelines by Cohen [29] were used to get an indication for the number of samples needed (rather than number of participants). Each sample was represented by a (state, action, next state, reward) tuple gathered from the conversation with Jamie. According to Cohen, for multiple regression analysis with 7 independent variables (two of the three variables for a person's state, and one variable for each of the five actions), a sample size of 102 is needed to observe a medium effect, for $\alpha = 0.05$ at power 0.80. Given that the number of possible persuasive strategies is 5, the final minimum sample size is $102 \cdot 5 = 510$ samples of (state, action, next state, reward) tuples. The number of samples gathered was 519.

Table 4.1: Participant characteristics from the pre-screening questionnaire

	Total (114)
Age	
Mean	29.56
Standard deviation	7.89
Range	20-56
Gender	
Man	56
Woman	54
Non-binary	4
Godin-Shephard leisure-time physical activity questionnaire results	
Insufficiently Active/Sedentary	98
Moderately Active	16
TTM stage for becoming more physically active	
Contemplating	60
Preparing	54
Average number of plans created in the past three months	
Plans for things that need to be done	7.31
Plans for walking	2.35
Enjoyment of creating plans	
Plans for things that need to be done	2.75
Plans for walking	1.46

4.1.4. Procedure

Pilot studies

A pilot study was run before the main study (between the 15th and 17th of June 2023), to ensure that all three steps of the study were working as intended. Based on the results of this pilot, bugs with saving some of the data to the database were fixed, and the measures which had diverging scales had these scales adjusted to (-10,10).

Another pilot study was run afterwards (between the 25th and 26th of June 2023), to ensure that all the data was gathered properly and that the new scales provided a meaningful change. The second pilot did not reveal any missed bugs or errors. As such, the data from this pilot was also included in the analysis. 24 samples from four people were provided by this pilot, and these are included in the statistics reported earlier. All participants in the pilots were excluded from the main study.

Study design

The study was observational, and consisted of three steps:

1. First, participants were informed about the study, and filled out the pre-screening questionnaire, which included a consent form. The pre-screening also gathered data about the person's situation in terms of confidence in following plans, perceived usefulness of planning activities, and attitude towards planning. The questionnaire took about five minutes to complete.
2. Participants who were eligible to participate in the interaction were then invited within one day of completing the pre-screening. Before the interaction, participants were informed about the virtual

coach and the goal of the interaction. They were told that they would have to communicate with the virtual coach through chat and that the purpose was to create a plan for taking walks. A brief explanation of why taking walks is a good way to become more physically active was also provided to the participants. The interaction took about 20 minutes.

3. After the interaction, participants who successfully completed the conversation were invited to fill in a post-questionnaire regarding their opinions of the virtual coach and of the conversation. This questionnaire took about three minutes.

The procedure for gathering the data is shown in Figure 4.1. In the first stage, people were excluded if they accessed the study from a mobile device. It was specified in the description of the study that the intended device to participate from was a desktop device. This was due to the fact that the conversational session needed plenty of space to display the chatbox and the plan table on the screen at the same time, which was not possible on mobile devices. After the week allotted to the study expired, there were still four people who did not complete the post-questionnaire. As such, the data provided by them in the post-questionnaire was not used, while the conversation data provided was included in the analysis.

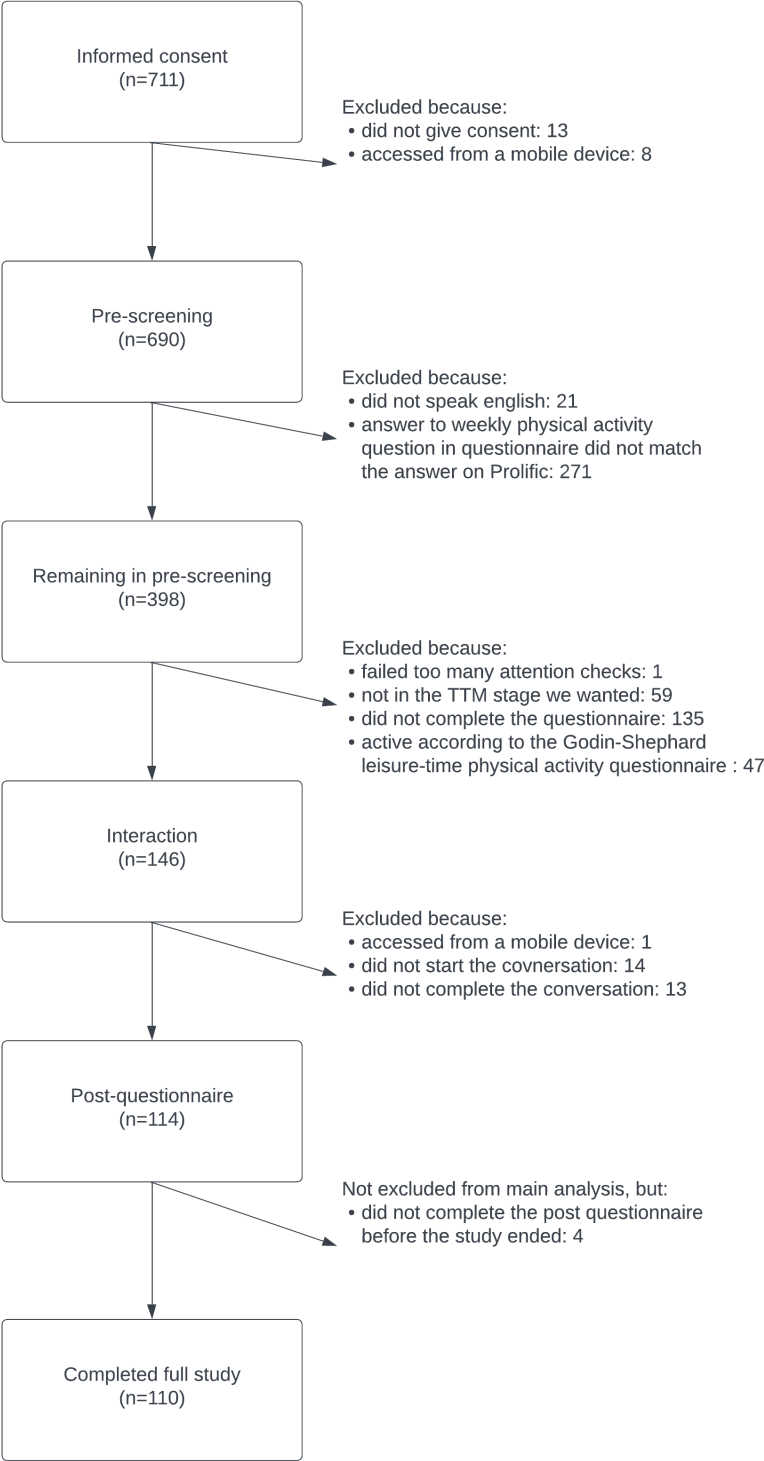


Figure 4.1: Overview of the stages of the study and exclusion criteria and participant flow

4.1.5. Data pre-processing and analysis

After gathering the data, it was first cleaned up in Python. Data from the people who were excluded from the analysis was removed and the remaining data was anonymised. In order to create the (state, action, next state, reward) tuples needed for the analysis, data from two separate tables in the database had to be combined to create the samples.

Additionally, there are state-action pairs for which there is little data (Appendix C). Since these are states which could occur in a real-life setting, they were kept as part of the model even if they appeared less frequently in the data. In order to have enough representative samples for each state-action pair, we imputed samples during the analysis. Imputing is done differently, depending on which type of state is considered in each question in the analysis.

The required number of samples was 510. If we consider states which **only contain the person's situation**, there are four possible states, and five possible actions per state, so 20 possible state-action combinations. Dividing 510 by 20 gives 25, meaning that, for state-action pairs with less than 25 samples, additional samples were imputed. For example, if there were 10 samples for a state-action pair, the resulting reward used in the analysis would be equal to $\frac{10}{25}$ of the original mean reward plus $\frac{15}{25}$ of the mean reward of all states, and the probability of transitioning to a next state would be equal to $\frac{10}{25}$ of the original probability of transitioning to said next state plus $\frac{15}{25}$ of $\frac{1}{4}$ (which is the probability of transitioning to a random next state). This kind of imputation was used when computing the optimal policy, since the alternative was to impute for 184 possible state-action pairs, meaning we would impute when there were less than three samples (since 510 divided by 184 is approximately equal to 3, which is not enough to provide representative samples for each of the states).

Similarly, if we consider **states which only contain what actions have been done**, there are 26 possible states, but not all five actions are possible in each state. In total, there are 63 possible state-action pairs, and dividing 510 by 63 gives 8, meaning that the mean values of state-action pairs with less than 8 samples had additional samples imputed during the analysis. The same example above can also be used in this case, by replacing the 25 with an 8.

After calculating these adjusted values for the state-action pairs which had too little samples, it was possible to answer the research sub-questions listed at the beginning of the chapter. Some adjustments regarding the analysis were made to the original OSF form. First, means which are outside of a credibility interval were interpreted as an indication of the values being different, as opposed to non-overlapping intervals. Second, the sixth question was changed from one investigating the order of actions depending on initial states to one comparing different policies, to be able to quantify the performance of the optimal policy.

Some of the questions required computing Q-values. This was done through value iteration, via the Bellman equation, where the discount factor was set to 0.85. Additionally, for some questions, 95% credibility intervals were computed, according to the guidelines provided by Oliphant [79].

4.2. Results

4.2.1. Q1. Which aspects of the person's situation are relevant to consider for the final model?

To select features for the reinforcement learning model, the G-algorithm [25] was used. The algorithm was modified, to accommodate features which have more than two possible values, through the two following changes. First, instead of using a t -test to get a p -value for significance, an analysis of variance (ANOVA) was used to obtain the p -value corresponding to picking a feature. Second, when blocks were split into sub-blocks by picking a feature, the next feature picked was the one that was the most relevant across all sub-blocks, rather than a possibly different feature per individual sub-block.

The G-algorithm was used to determine which parts of a person's situation were relevant for the optimal policy (Q-values). Q-values were used to pick which features to select, since they are the core of the reinforcement learning model and the basis for forming the policies.

Additionally, the algorithm could select between two, three, and four values for each of the features. For the case where the algorithm checked if the feature should have two values, the data was split at the 50th percentile, resulting in two subsets (0th to 50th percentile and 50th to 100th percentile), with similar splits at the 33rd and 66th percentiles when three values were checked, and splits at the 25th, 50th, and 75th percentiles, when four possible values were checked.

Only two features were selected, since the objective was to reduce the state space by eliminating

one of the three features which make up the person's situation. These features were confidence, which was allowed to have two values, and perceived usefulness which also had two possible values. The process for selecting the features and values is explained in detail in Appendix D.

Next, the data was restructured, to allow the states to reflect the features and values picked. In the following questions, states only contain confidence with two values and perceived usefulness with two values as part of the person's situation.

4.2.2. Q2. How well can states at the end of the dialogue predict their corresponding rewards?

If states at the end of the dialogue give an accurate prediction of their corresponding reward, then it means those states are a good indication of the person's commitment to the plan. To check if states can predict the rewards, we used leave-one-out cross-validation, leaving out end states, to check how well a model could predict the reward of the end state using two approaches:

1. taking the mean reward of all end states and using it to predict the reward of the left-out sample,
2. taking the mean reward of all end states that were the same as the left-out sample and using this mean to predict the reward of the left-out sample.

For both of these approaches, the mean L1 error was computed. The approach with the lower error was considered to be better for predicting the reward, since a lower error means a better prediction. Additionally, for each state, the mean reward at the end of the dialogue is computed. The results for this question can be seen in Figure 4.2, where only the person's situation is displayed for each state. The bars show the mean probabilities of the prediction being correct, while the black lines represent the 95% credibility intervals. The solid line is the mean reward over all of the end states, while the dashed line is the mean reward of the end states per state. In terms of mean rewards, the state ['0', '0'], corresponding to low confidence and low perceived usefulness had the lowest reward at the end of the dialogue, while the state ['1', '1'], corresponding to high confidence and high perceived usefulness is the one with the highest reward. The state ['0', '1'] also has a reward that is above the overall mean, so high perceived usefulness seems to correspond to a high reward.

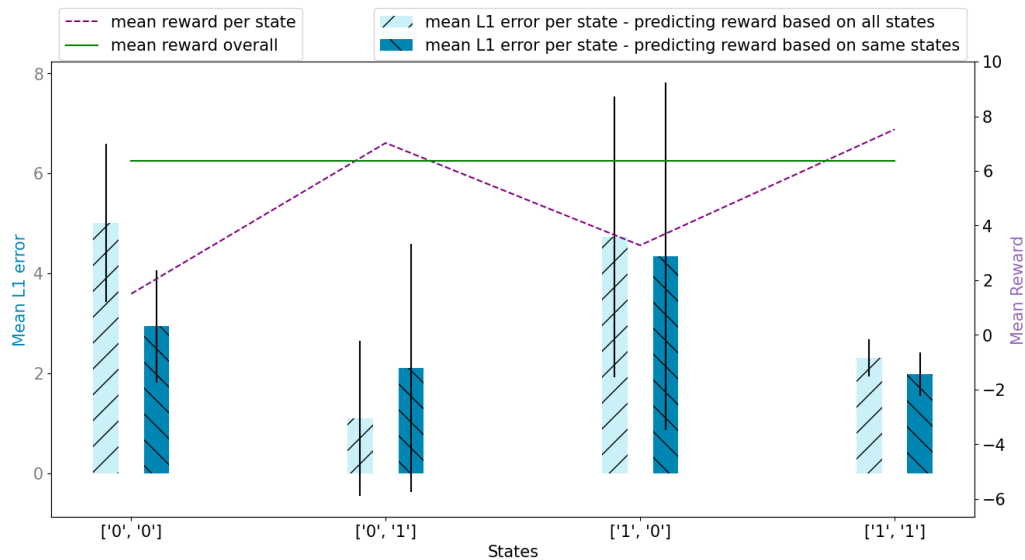


Figure 4.2: The mean L1 errors for the two approaches and the mean rewards per state. States are of the form ['confidence', 'perceived usefulness']. For both confidence and perceived usefulness, '0' means low confidence or perceived usefulness, while '1' means high confidence or perceived usefulness. The axis on the left shows the mean L1 error, while the axis on the right shows the mean reward.

From the means which overlap the credibility intervals, it is noticeable that the two approaches give similar results, except for the state ['0', '0']. This is likely due to the fact that this state's reward is far from the mean reward of all the states, which means that predicting that its rewards are equal to the

mean reward of all states performs worse than for the other states. It is sometimes the case that there are only a few samples which are the same as the state that is left out, causing some error bars to have large 95% credibility intervals.

4.2.3. Q3. How well can states based on the person's situation predict states after a persuasive strategy used by the conversational agent?

Being able to predict states after a persuasive attempt could give an indication of whether or not the dialogue is progressing towards a state in which the person is likely to commit to the plan. For this question, leave-one-out cross-validation was used, leaving out all samples from the same person when computing transitions (since these samples are not independent from each other), to compare:

1. predicting the next state using the transition function learned from the training data,
2. predicting that the next state stays the same,
3. predicting the next state while having an equal probability for each next state.

The probability of moving to the next state was used to compare the approaches. A higher probability means the prediction was better for that approach. The results for this question can be seen in Figure 4.3. The bars show the mean probabilities of the prediction being correct, while the black lines represent the 95% credibility intervals. In all of the cases, predicting that the state stays the same has the highest probability of being correct. All means of an approach are outside of the credibility interval of the other approaches. Giving an equal probability to each next state is outperformed in all cases, showing that either using the transition function or predicting that the state stays the same is much more likely to result in a correct prediction, both of which are correct more than 50% of the time.

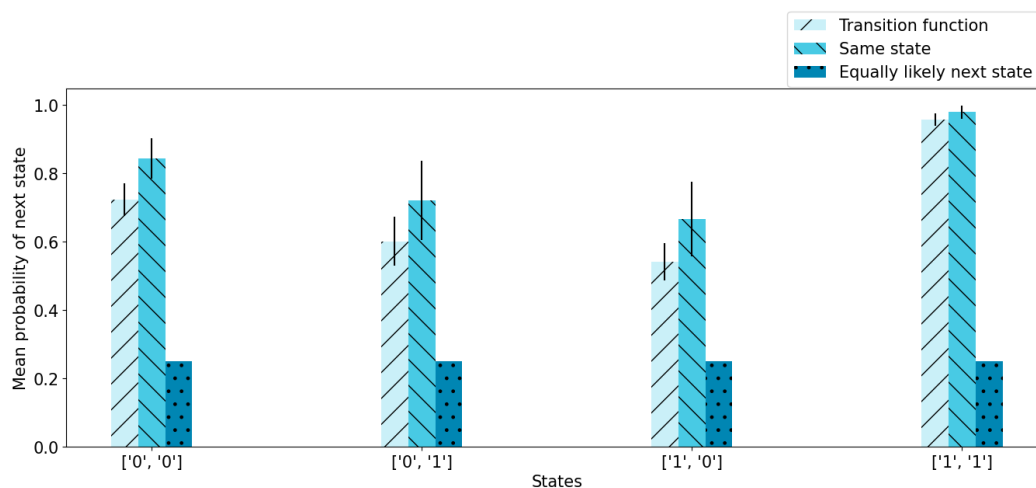


Figure 4.3: The probability of correctly predicting the next state based on the current state for each of the three approaches. States are of the form [‘confidence’, ‘perceived usefulness’]

4.2.4. Q4. What is the optimal model for persuading people to follow plans for walking?

From Q1, it was determined that the person’s situation would be composed on confidence with two values and perceived usefulness with two values in the final model.

The optimal policy was computed via value iteration, and is presented in its entirety in Appendix E. While the first change to the plan happens early on during the discussion, the second one is usually reserved for the fifth time step, or sometimes the sixth time step, but never earlier than that. This might indicate that one change at the beginning lets people familiarise themselves with the plan, and after going through the dialogue and learning more about possible barriers or strategies for accomplishing plans, a second opportunity to change is optimal.

Additionally, there are states for which the best policy is to pick, with equal probability, a persuasive strategy from the ones which are possible in that specific state. This is caused by the fact that there are states for which there is little data, so no best action can be picked for those states. In Appendix C, a distribution of how many times each action was done in each of the states is shown.

4.2.5. Q5. What is the effect of (multiple) optimal persuasive strategies on persuadees' states?

Ideally, we want multiple successive persuasive strategies to change a person's state to the one with high confidence and perceived usefulness, since then they might be more likely commit to following the plan.

A simulation of how people would transition between states was run, following the optimal policy from Q4 and the transition function learned from the data. The simulation started with a uniform distribution of 100 people in each of the four possible starting states. At each time step, the best action for a specific state was selected, according to the optimal policy. Then, by using the transition function, people transitioned to a new state and then the simulation continued for the following time step. The simulation ran for 6 time steps in total, which is the maximum amount of time steps it could have in a real setting (corresponding to the virtual coach using up all the persuasive strategies).

Through this simulation, we checked:

1. what percentage of people reached the "good" state (the maximum value for each of the features - the state ['1', '1']) at the end of their dialogue,
2. how many time steps on average it took people to reach the "good" state and end the conversation.

Figure 4.4 shows the distribution of people after simulating six time steps of following the optimal policy. The bars show what percentage of people were in a specific state at the start of the simulation (time step zero), and after each persuasive strategy (time steps one through six). The bars corresponding to the label "Ended" indicate the cumulative percentage of people that ended the conversation by reaching the "good" state, where both features are high. By the end of the sixth time step, this percentage is about 82%. On average, it takes approximately 3 persuasive strategies to bring a person to the good state and end the conversation.

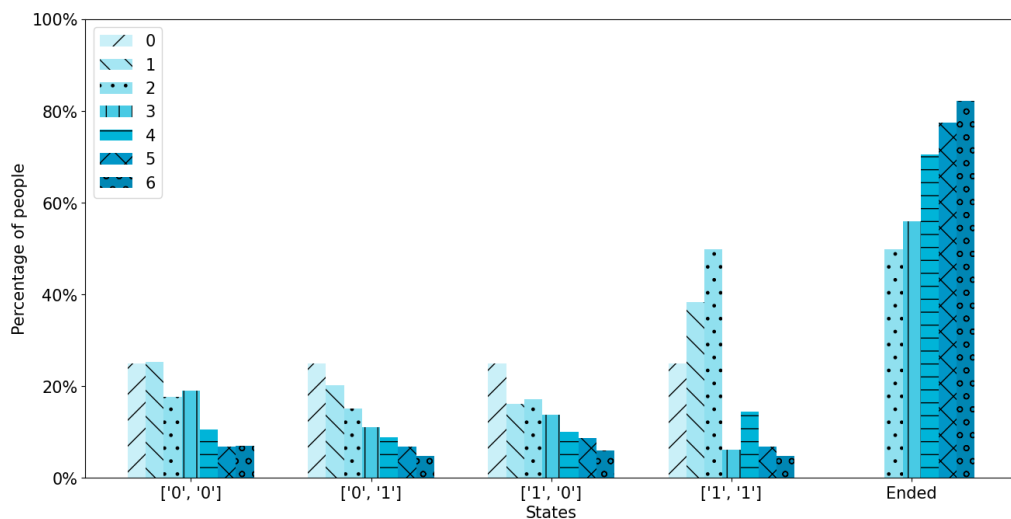


Figure 4.4: The distribution across states after six time steps for the optimal policy. States are of the form ['confidence', 'perceived usefulness']

Given the probabilities above 50% of staying in the same state (seen in Q3), it would seem more likely that people stay in the same state than move to a different one. However, over multiple time steps people tend to move to the state with high confidence and perceived usefulness, which is the

one where they end the conversation. For all states other than the aforementioned one, the percentage of people in each state tends to decrease over time. This means that people tend to move to the state with high confidence and perceived usefulness and end the conversation.

4.2.6. Q6. How do different policies perform, compared to the optimal policy?

Seeing how the optimal policy performs compared to other policies could give an indication of whether following the optimal policy is better (or possibly worse) than other non-optimal policies.

Besides the optimal policy (π^*), two other policies are computed: the worst policy (W, shown in Appendix F) and the optimal policy obtained by removing the person's situation from the state variables (NO_S shown in Appendix G). W is obtained by taking the worst possible action in each case. If it is similar to π^* , then this could indicate that the persuasive strategies have similar effects. NO_S indicates a fixed order of actions, as explained in appendix G. If NO_S performs better than π^* , it means that considering the person's situation is not relevant, since the persuasive strategies would be used in the same order, regardless of the person's situation.

First, we looked at how these two new policies transition people from one state to another. To do so, simulations using the same setup as in Q5 were ran for these policies. The results can be seen in Figures 4.5 and 4.6. The simulations seem to indicate that the three policies are similar to each other in terms of how people transition from one state to another: despite the high probability of a state remaining the same, over multiple actions, more than 70% of people transition to the state with high confidence and perceived usefulness after six persuasive strategies are used. NO_S also takes 3 actions on average to transition people to the state with high confidence and perceived usefulness, while it takes 4 actions on average for W to do so.

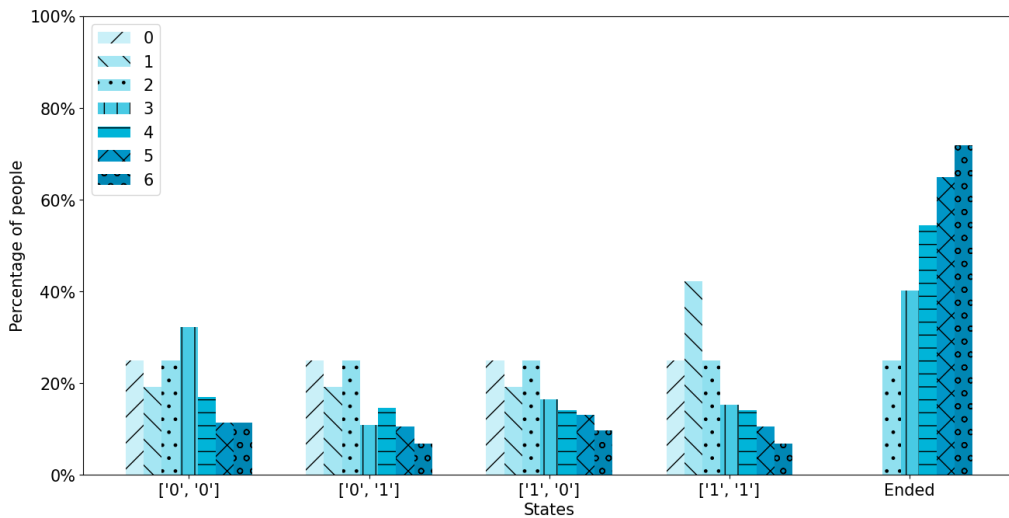


Figure 4.5: The distribution across states after 6 time steps for the worst policy. States are of the form ['confidence', 'perceived usefulness'].

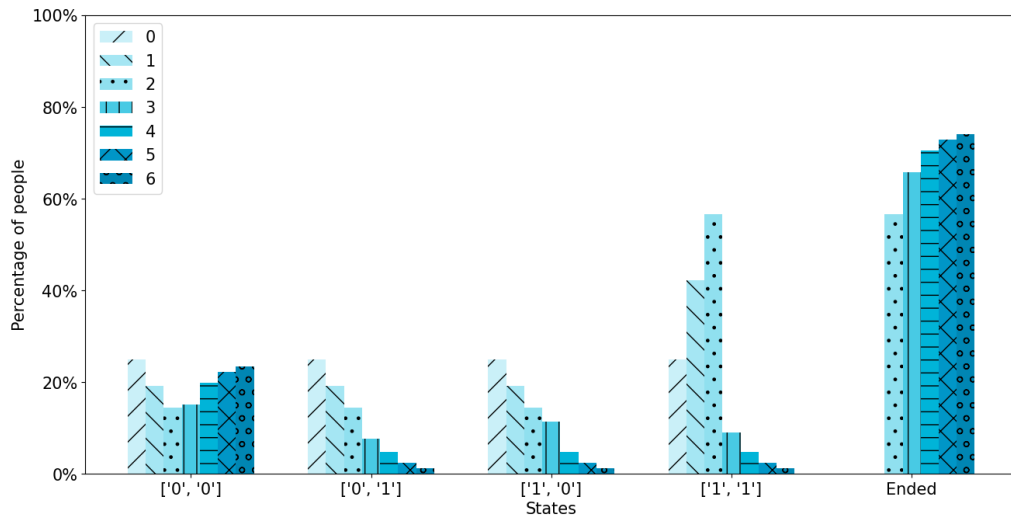


Figure 4.6: The distribution across states after 6 time steps for the policy without the person's situation. States are of the form ['confidence', 'perceived usefulness'].

To further investigate the differences between the three policies, another set of simulations was run for all three policies. The difference was that, instead of starting the simulation with 100 people in each starting state, this number was set to the number of samples available per state. For example, if there were 5 samples of people starting in state [0, '0', '0', False, False, False, False] (where the first number represents the number of modifications done to the plan, the second and third numbers are the confidence and perceived usefulness which we have been using, and the last four boolean values indicate if the corresponding action was done or not), then the simulation would only have 5 people in this state, rather than 100.

The entire dialogue was simulated, then the mean reward of each time step was computed. The goal was to have a dialogue that is as short as possible, so the rewards after the second time step (which is the earliest possible end for a dialogue) were discounted, and the discount would increase for each subsequent time step. Meaning that, if the discount factor was 0.85, for example, the mean reward in time step two would not change, the mean reward in time step three would be multiplied by 0.85, the mean reward in time step four would be multiplied by 0.85 squared, and so on. After this was done, the mean reward of all time steps for a policy was computed. Several different discount factors were tried and the results are shown in Figure 4.7. For all discount factors, π^* has a slightly larger discounted mean reward than NO_S, and W has the smallest discounted mean reward. While this is the case, the difference between the three policies are rather small, as can be seen from the lines which are close together in the figure and the overlapping 95% credibility intervals.

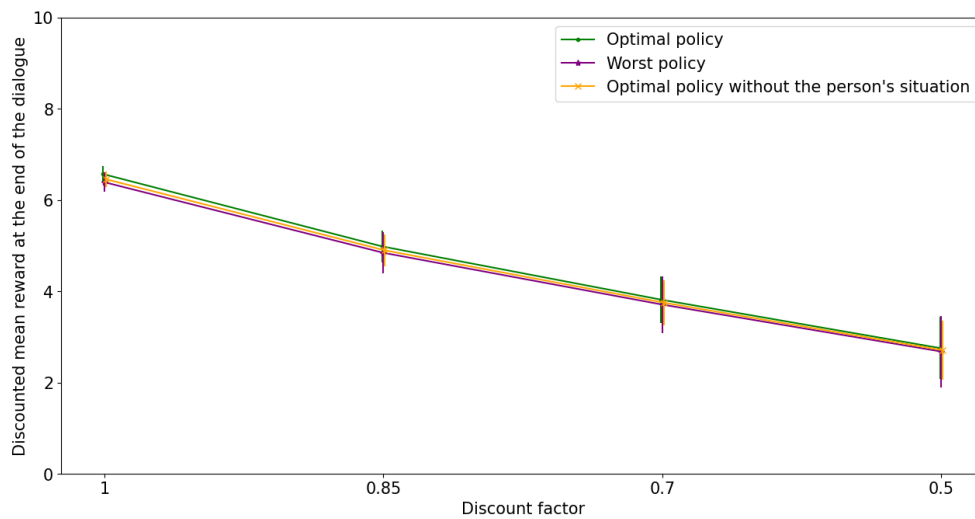


Figure 4.7: The discounted mean rewards of each policy.

4.2.7. Exploratory analysis

Effects of the persuasive strategy chosen

The results of Q6 pointed towards the possibility that different persuasive strategies have a similar effect in the long term, due to the similarity between the optimal and worst policies. Thus, the importance of the persuasive strategy chosen in each state is further investigated. Considering the two features chosen for the person's situation - confidence and perceived usefulness - we can check if the use of a persuasive strategy changes this feature, by using a paired Bayesian t -test to compare the feature before and after using the persuasive strategy, and also quantify the size of this difference by checking Cohen's d .

The results, explained in Appendix H, indicate that individual persuasive strategies have a small effect on the state, while multiple persuasive strategies over the course of the entire dialogue show a moderate effect.

Participants' opinions and frequency of plan creation

Using the data from the pre-screening questionnaire, it was possible to determine if people usually create plans, and if they enjoy doing so. If someone does not enjoy creating plans, then it is likely that this would impact the reward at the end of the dialogue (the person would probably not be satisfied with a dialogue regarding planning, and might not want to commit to following the plan).

The participants specified that they created an average of 7.31 plans for things they need to do in the future, and an average of 2.35 plans for walking during the past three months. In terms of their enjoyment of creating plans for things they need to do in the future, people said it makes them feel moderately good (2.75 average on a -5 to 5 scale). For plans for taking walks, people specified that it makes them feel slightly good (1.46 average on a -5 to 5 scale).

Thematic analysis of motivating and demotivating aspects of the interaction

In terms of opinions regarding the dialogue, people were asked in the post questionnaire about what they found motivating and what they found demotivating about the interaction with the virtual coach. The thematic analysis followed the guidelines provided by Braun and Clarke [15] and is described in Appendix I.

Four common themes were identified for the motivating and demotivating aspects of the interaction: the persuasive strategies, the planning, the virtual coach, and not finding anything motivating or demotivating. In terms of persuasive strategies, people appreciated coping planning (identifying and dealing with barriers), but the testimonials were sometimes seen as motivating and other times as demotivating. Having a plan that seems achievable and that takes into account when a person had free time was motivating for people, but the lack of being able to see and customise the plan after the first two weeks

was disappointing. While the virtual coach's messages were seen as encouraging, people disliked that it seemed robotic in its answers and that it repeated questions many times, making the coach seem talkative due to the long dialogue. Finally, some people indicated that nothing was motivating, while other people indicated that nothing was demotivating, which forms the last theme.

Answers to the ASA questionnaire

Questions regarding the agent are also asked in the post questionnaire. The artificial social agents questionnaire helped us determine people's opinions regarding the coach, such as believability, usability, performance, and opinions regarding the interaction, such as user acceptance, user engagement, and user-agent interplay. Based on 110 responses to the post-questionnaire, Jamie's average score for the ASA questionnaire is 10. Other disembodied agents which communicate through spoken language have similar scores to Jamie (even though Jamie communicates through written language). These include Apple's Siri, which scored a 13, and the fictional character from "A Space Odyssey" named HAL 9000, which scored a 14 [39]. The ASA chart for the comparisons between Jamie, Siri, HAL 9000 can be seen in Figure 4.8. The chart was generated using the (slightly adjusted) ASA chart tool [37], with the means for the other ASAs from Fitrianie et al. [38].

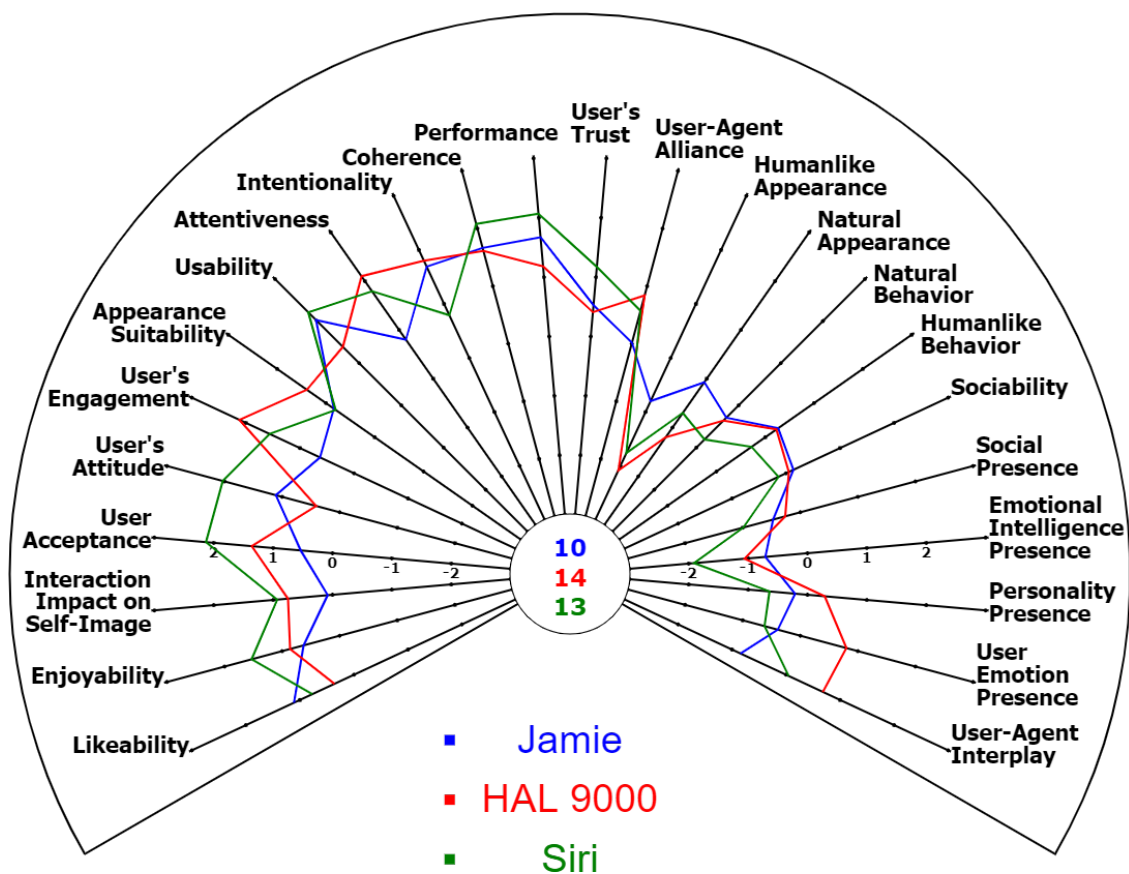


Figure 4.8: The ASA chart comparing Jamie, HAL 9000, and Siri.

Compared to the average scores of Siri and HAL 9000, Jamie was more liked by people, its appearance was more humanlike and natural, and its emotional intelligence presence is higher. The results also show similarities with the two other ASAs, in usability, performance, and coherence. However, there are also items for which Jamie scored lower than the other ASAs, such as its attentiveness, enjoyability, user acceptance, and user engagement.

4.3. Discussion of the results

By answering the main research questions, the model for persuading people to commit to following plans for taking walks was presented.

As a first step, we used the G-algorithm to select confidence with two values and perceived usefulness with two values as part of the final model. Thus the state space was reduced from three features in the person's situation to two.

By investigating the rewards at the end of the dialogue, we learned that the states at the end of the dialogue give some indication of what the reward will be. However, predicting the reward based on the mean reward overall or on the mean rewards of the states which are the same are not very different from each other. Since the state with high confidence and perceived usefulness has the largest mean reward, it seems to be the case that people who reach this state are the most likely to commit to following the plan.

The transition function learned from the data allows states to predict future states after a persuasive attempt with an accuracy which is better than picking a random state with equal probability. At the same time, predicting that the state does not change has an even better probability of being correct. This means that, on average, people are not very likely to move to a different state after a single persuasive strategy is used. However, based on the results of the simulations, we have shown that following the optimal policy and using multiple persuasive strategies can bring people to a state in which they are likely to commit to following the plans. The optimal policy focuses on getting people to the state with high confidence and perceived usefulness as early as possible, since then they can end the dialogue.

When looking at the optimal policy learned from the data, we observed that the action of making changes to the plan was done once early on in the conversation and once late in the conversation. This indicates that giving people some time to process the persuasive strategies of the virtual coach before changing the plan again is optimal.

The comparison of different policies showed that the optimal policy is similar to the worst policy and to the optimal policy which disregards the person's situation (which is equivalent to a fixed order of persuasive strategies), in terms of the rewards obtained at the end of the dialogue. When looking at the percentage of people who end the dialogue in the state with high confidence and perceived usefulness, the optimal policy which considers the person's situation performed better than the optimal policy which does not, by approximately 8%. This might indicate that including the person's situation in the model is beneficial, since it achieved better performance.

While including the person's situation gets better results than not including it, both approaches have their disadvantages. Determining the person's situation and how it changes requires asking the same questions multiple times, and thematic analysis of demotivating aspects of the conversation revealed that people disliked answering these repeated questions. Nagata, Hashiguchi, and Sadoun [74] showed that a conversational agent that repeats the same question is perceived as annoying and might be demotivating for the person it interacts with, which is in line with our results. This constitutes the disadvantage of including the person's situation, and given the issues with repeating questions, it might be better to use the other policy, to avoid demotivating people. For the policy which does not use the person's situation, the disadvantage is that removing the person's situation means that the agent no longer has a way of determining when the person is likely to commit to following the plan. This means that a conversation which could have ended after, for example, three persuasive strategies will continue until all persuasive strategies are used up.

A possible solution to both of these disadvantages is to learn when to ask the questions about the person's situation. This is often done in conversational recommender systems, whose purpose is to make recommendations based on people's preferences. For example, Tian et al. [98] present a recommender system which uses deep Q-learning to determine when their agent should ask the person a question about their preferences and when it should make the recommendation. In our setting, we could learn when to ask the person about their situation, which would avoid asking too many times. Additionally, if the person's is in a state where with high confidence and perceived usefulness, then the virtual coach would know it should end the conversation.

Further exploratory analysis showed that individual persuasive strategies only have small effects on the person's situation. However, when comparing states at the beginning and end of the conversations, we found a moderate effect. It is possible that time influences how much of an effect the persuasive strategy has. The sleeper effect in persuasion suggests that information from a source which is not seen as credible first has a small impact, but, if the source is then seen as reliable, the impact increases

[51, 54]. Thus, it is possible that people initially saw the virtual coach as unreliable, but, after going through the dialogue and reading the information that the coach was providing, their opinion changed and so the persuasive strategies started having a more meaningful effect in the end. Additionally, the elaboration likelihood model [80] suggests that in-depth processing is necessary to see change. Then, since in-depth processing takes some time, changes cannot be observed in the short term, which is also what we observed. By the end of the interaction (which took about 20 minutes on average), people might have processed the persuasive strategies in more depth and then a change was visible.

Another possible explanation for the small effects of the persuasive strategies is related to the transtheoretical model stage that people were in. Ferron and Massa [35] indicate that people in different stages require different types of motivation. For example, people in the preparation stage need extrinsic motivation in the form of external regulation (e.g. people wanting to show others that they are in good shape). The people in our study were in the contemplation or preparation stages, so they could benefit from a gamified system with badges and points which allows them to show off their progress to their peers, and which can increase engagement [65]. Additionally, the authors mention that, for people in one of these two stages, increasing the importance of the cognitive dissonance between people's intention to be active and the lack of actually doing so can give them the push to start changing their behaviour. The action of explaining why planning is useful did have some education elements, but they were not focused on increasing the importance of the dissonance, which might have had a larger effect.

Lastly, based on the answers to the short form ASA questionnaire, we compared Jamie to two other ASAs which communicate in a similar manner: Siri and HAL 9000. People indicated that Jamie's appearance was more natural and humanlike, but also that it was less attentive and engaging. In terms of usability, performance, and coherence, all three ASAs were similar.

4.3.1. Limitations

The main limitation of this study is the relatively low amount of data. Considering the size of the state space and the method of gathering the data about the states, it was difficult to collect many samples given the duration of the study. The issue with the low amount of data is the possibility that the optimal policy learned from the samples gathered does not generalise to a larger population. Although we tried to balance participants across all possible initial states to collect enough data to provide a representative sample, it was ultimately impossible to distribute them evenly. Some initial states had more participants, while others had less (Appendix C). This, combined with the impossibility of collecting a large amount of data, lead to having states where the virtual coach would need to pick a random persuasive strategy. To compensate for the lack of data, for state-action pairs which did not have enough data, we imputed average data. This means that the results of the analysis might not be fully representative of the data gathered, further reinforcing the need for collecting more data to reach solid conclusions. Still, the sample size established based on the guidelines by Cohen [28] was reached, and simulated results show that the agent is effective in bringing people to the state with high confidence and perceived usefulness, where they are most likely to commit to following the plan.

The people who participated in the study indicated that they somewhat enjoy creating plans in the pre-screening questionnaire. As such, the sample might not be representative of people who do not enjoy creating plans at all. Furthermore, even though they specified a moderate degree of enjoyment for creating plans for taking walks, they only created 2.4 plans on average in the previous three months. Thus, they had little recent experience with creating a plan for walking, which could mean that the novelty of this action was why they found the action and coping planning to be motivating aspects of the interaction.

Another limitation comes from the fact that all of the data for the model was provided through subjective measures, through the questions that were asked. As the study is aimed at people who were inactive, it is possible that the effects of the social desirability bias [17] played a role in how these people answered the questions. Participants could have provided answers which would make them appear willing to be more physically active, even if this was not necessarily the case.

A final limitation is given by the mode in which answers were recorded. Since people had to answer the same question multiple times, it is possible that they stopped considering the question and the provided options in their entirety and just picked the same answer they had used previously. Indeed, for some participants it was the case that their answers were very similar or even identical, even after all the persuasive strategies had been used. However, attention checks in the pre-screening questionnaire and the detailed answers to open-ended questions during the conversation seem to indicate that people

were indeed answering truthfully and that their states did not change much after individual persuasive strategies.

5

Conclusion and Discussion

This chapter provides a discussion and conclusion of the Masters thesis. The research questions are restated and answered, then the contributions to the field are presented. Limitations and suggestions for future work are listed at the end, along with closing remarks.

5.1. Conclusions

The research question answered in this study was:

How can reinforcement learning be used to persuade people to follow their dyadic physical activity plan in a dialogue with a virtual coach?

This question was further split into sub-questions, for which answers are provided below.

5.1.1. What are factors, values, and concerns that need to be considered in a dialogue that facilitates goal setting and dyadic planning with a user?

To answer this question, literature and experts were consulted, resulting in a list of factors, values, and concerns for dialogues which aim to develop a dyadic plan.

Starting from the theory of setting goals, we found that plans can help people commit to their goals and that dyadic plans are more effective at doing so than individual plans. Additionally, the goals should be personalised to the person and, since the people in our study are inactive, the most appropriate types of goals are walking goals. Since the dialogue should make people commit to doing physical activity according to the plan, action and coping planning are powerful tools which can help guide people towards their goals.

When creating a plan, it is important that people are given the ability to create the plan themselves, rather than be assigned a plan. Creating the plan themselves can also help them remember and enact the behaviour defined in the plan, thus making progress towards their goal. However, people should also be guided through the process, since they might be overconfident in their abilities if they do not have much experience with physical activity or might underestimate how much time it takes to do an activity. As such, the virtual coach should be able to provide assistance and relevant information for creating plans. The virtual coach should ask people about routines, anticipated energy levels, and past experience with physical activity to help them create a plan, and compile this information into a plan that is easy to interpret and that does not stretch too far out into the future, since planning too far into the future is difficult and can be overwhelming. Adapting to the person was found useful in changing people's behaviour, and a viable technique for doing so is reinforcement learning.

5.1.2. How can a dialogue for dyadic planning with a virtual coach be designed?

We identified two major goals that a dialogue for dyadic planning should have: plan creation and plan commitment.

For plan creation, aspects such as routines and energy levels were found to be relevant for creating plans that people commit to. Thus, part of the dialogue included asking people when they are free and

when they have more energy than usual. The next part of the dialogue was setting a goal and creating a plan for reaching this goal, based on free time and energy levels.

The plan commitment aspect of the conversation was approached as a reinforcement learning problem. For this, we designed the core part of the dialogue as a loop where the person could specify their situation, consisting of aspects which were found to be relevant to approach when persuading someone to follow a plan: confidence, perceived usefulness, and attitude. Following this, the virtual coach would use one of its persuasive strategies, and then ask the person to specify these three aspects again. The persuasive strategies selected were: making changes to the plan, explaining why planning is useful, coping planning (though identifying and dealing with barriers), and showing testimonials.

The dialogue ended with a measurement of the reward signal, which was composed of: satisfaction, confidence in reaching the goal, and commitment to the plan. Commitment to the plan was interpreted as a concept representative of a person's commitment to the goal, which is an important aspect of goal-setting.

5.1.3. How well can reinforcement learning select persuasive strategies to be used as part of a dyadic planning dialogue to make people satisfied with the dialogue, committed to the plan for taking walks, and confident in reaching their goal?

To answer this question, data was gathered through an observational study, and a reinforcement learning model was implemented, based on the data. The final model's states were composed of the person's situation and the persuasive strategies that were already used by the virtual coach. Two aspects of a person's situation were selected to be part of the final model: their confidence in following the plan for walking, and their perceived usefulness of planning for physical activity. These states were able to give some indication of the reward, meaning that they could somewhat predict if people were likely to commit to the plan.

While investigating the way in which people move from one state to another, we noticed that people had a tendency to remain in the same state (meaning that their confidence and perceived usefulness would not change most of the time). This hinted at the possibility that different persuasive strategies have similar effects on a person's situation.

A Bayesian t -test comparing the person's situation before and after each persuasive strategy showed that there is no difference in the situation. The effects of each individual action on a person's state were mostly small (and, in some cases, moderate), according to Cohen's d . At the same time, for multiple successive persuasive strategies there was a difference, and this difference was moderate, according to Cohen's d .

According to the elaboration likelihood model, in-depth processing is needed to observe a change. As there was little time between different persuasive strategies, we hypothesise that in-depth processing of individual persuasive strategies was not achieved. At the same time, since the dialogue took about 20 minutes on average, the interaction as a whole could have been processed in-depth, thus producing the moderate effect of multiple successive persuasive strategies. Furthermore, if people initially saw the virtual coach as unreliable, but the information it gave during the dialogue caused them to change their opinion, the sleeper effect [54] could explain the moderate impact of the entire dialogue. Additionally, it could have been the case that the person's transtheoretical model stage should have been taken into account, since different stages require different types of motivation to increase their engagement.

To investigate the optimal policy further, we compared the optimal policy (obtained by selecting the best persuasive strategy for each state) with two other policies: the worst policy, obtained by selecting the worst possible persuasive action in each state, and an optimal policy where states did not include the person's situation. The results of additional simulations showed that all three policies had similar results in terms of the rewards at the end of the dialogue. Regarding the percentage of people that reach a state in which they are likely to commit to following their plan the results were: 72% for the worst policy, 74% for the optimal policy that does not consider a person's situation, and 82% for the optimal policy which does consider a person's situation. Thus, there were some differences between the policies, indicating that the policy which considers a person's situation performs better than the one that does not. At the same time, the latter has some advantages, such as not having to ask people to specify their situation multiple times during the dialogue, which was found to be demotivating by people.

Thus, there seems to be a trade-off between the performance of the virtual coach and the motivation of the people taking part in the interaction. So, while the policy which includes the person's situation achieves results which are 8% higher, it might be more appropriate to value personal motivation higher than performance and opt for not including the person's situation in the state, which is equivalent to following a fixed order of persuasive strategies.

5.2. Contributions

The first contribution of this thesis is the list of factors, values, and concerns for dyadic dialogues with virtual coaches, based on recent action planning research and its relation to goal-setting theory.

The second contribution is the design, development, and evaluation of a reinforcement learning model for a virtual coach that can help people become more physically active through dyadic planning.

A third contribution is the data gathered during the study. The dataset of 519 samples of (state, action, next state, reward) tuples can be used to develop different models, based on different combinations of the state components or of the reward function, which might work in different contexts. The free text responses to prompts from the virtual coach can also serve as a basis for future research into how people perceived planning and what they found important, or into the ways in which people created a coping plan for physical activity. The dataset includes opinions regarding the usefulness of planning, personal descriptions of common barriers to physical activity, as well as ways to address those barriers, and thoughts about the important takeaways of testimonials from other people who created plans for physical activity.

The fourth contribution is the analysis of the developed model, which can serve as a starting point for further research into dyadic planning with a virtual coach. The analysis provides the optimal policy to be used when trying to persuade people, the transition function learned from the data, and templates for running simulations of the model. Other combinations of features for the states can be considered, possibly leading to new insights regarding how the optimal policy that includes the person's situation compares to the optimal policy that does not.

5.3. Limitations

The main limitation of this work is not having tested the final model, which might be a threat to external validity. Thus, while the model provided is one of the main contributions, there is no quantifiable measure of how well the model would perform in a real-life setting. Still, results of the analysis and the simulations show that the model has the capacity to persuade people to commit to following the plan, while open ended answers from the post questionnaire suggests that people felt encouraged to be more physically active and seemed to enjoy creating the action and coping plans, even if creating plans is not something they usually do (based on results of the pre-screening questionnaire).

Another limitation comes from only having interacted with two experts in the health domain, and only having consulted literature regarding goal-setting and planning for physical activity. Therefore, our view of planning in other domains might be too narrow to allow the virtual coach to seamlessly operate in different domains which might consider planning and goal-setting, such as self-management planning, organisational planning, or students planning in an academic setting.

A third limitation is the type of goals considered. Literature [88] indicated that walking goals are suitable for people who are inactive, but personal preferences could have also played a role in how people perceived the dialogue with the virtual coach. Other types of physical activity could have been more enjoyable for some people, meaning that their satisfaction with the interaction could have been impacted by this.

5.4. Future work

A first avenue for future work is to test the performance of the model in a real-life setting. An experiment could provide insight into how effective the optimal policy is, and clearly justify whether the virtual coach achieves its objective of persuading people to follow plans for taking walks. Given that simulations have already been run, the long-term effect of the interaction with the virtual coach could be studied. It would be interesting to check whether people really commit to following the plan in the long term or if the commitment is only temporary and fades as soon as effort is necessary.

Both the structure of the conversation and the persuasive strategies could serve as a basis for a goal-setting dialogue which includes planning. The persuasive strategies can be replaced by different strategies, while the structure of the dialogue could be altered to include more varied considerations. For example, the similarity measures used in Albers et al. [5] could further help determine which action is suitable, based on what actions were suitable for similar people in the past. We have seen that some people indicated that the testimonials shown to them were not relevant because they came from someone very different from themselves. Thus, considering how similar people are in terms of how often they exercise, their age, and their sex, as was done by Albers et al. [5], could provide better results when trying to persuade people based on testimonials.

Furthermore, the design of a virtual coach that helps people create plans for physical activity can be expanded upon, to create new models which could incorporate more aspects which are relevant to determine what the best persuasive strategy should be. Factors such as internal motivation through internal rewards [63] or self-efficacy [8] could be used to gain a better representation of a person's situation, but the problem of asking people the same questions multiple times still remains.

As such, better ways of measuring the aspects of a person's situation could also be investigated. In the study performed, many people complained about having to answer the same questions repeatedly, so a different method for gathering this data might be more suitable. For example, Calvo et al. [23] present different ways in which natural language processing was used to determine emotions from text. Through free text answers, an NLP model could infer a person's state without having to directly ask for the specific measures. This could also improve how natural people perceive the conversation to be, by eliminating repetition and breaking of the conversation flow when asking questions about the state.

5.5. Closing remarks

Through this thesis, we created a model for a virtual coach who persuaded people to follow plans for walking created through dyadic planning. The reinforcement learning model was designed based on factors, values, and concerns identified in action planning research, and developed based on data gathered in an observational study. The analysis shows that the appearance and behaviour of the virtual coach are suitable in a planning context. Furthermore, we have shown that taking into account the situation that a person is in performs better than using persuasive strategies in a fixed order.

References

- [1] Marc A Adams et al. “An adaptive physical activity intervention for overweight adults: a randomized controlled trial”. In: *PloS one* 8.12 (2013), e82901.
- [2] Rangina Ahmad, Dominik Siemon, and Susanne Robra-Bissantz. “Communicating with Machines: Conversational Agents with Personality and the Role of Extraversion”. In: Jan. 2021. DOI: 10.24251/HICSS.2021.492.
- [3] Icek Ajzen. “The theory of planned behavior”. In: *Organizational behavior and human decision processes* 50.2 (1991), pp. 179–211.
- [4] Nele Albers, Mark A Neerincx, and Willem-Paul Brinkman. “Addressing people’s current and future states in a reinforcement learning algorithm for persuading to quit smoking and to be physically active”. In: *Plos one* 17.12 (2022), e0277295.
- [5] Nele Albers et al. “Setting physical activity goals with a virtual coach: vicarious experiences, personalization and acceptance”. In: *Journal of Medical Systems* 47.1 (2023), p. 15.
- [6] Sandra Andraszewicz et al. “An introduction to Bayesian hypothesis testing for management research”. In: *Journal of Management* 41.2 (2015), pp. 521–543.
- [7] Aurélie Baillot et al. “Physical activity motives, barriers, and preferences in people with obesity: A systematic review”. In: *PloS one* 16.6 (2021), e0253114.
- [8] Albert Bandura. “Social cognitive theory of self-regulation”. In: *Organizational behavior and human decision processes* 50.2 (1991), pp. 248–287.
- [9] Amy Bantham et al. “Overcoming barriers to physical activity in underserved populations”. In: *Progress in cardiovascular diseases* 64 (2021), pp. 64–71.
- [10] Richard Bellman. “Dynamic programming”. In: *Science* 153.3731 (1966), pp. 34–37.
- [11] Corina Berli et al. “Interpersonal processes of couples’ daily support for goal pursuit: The example of physical activity”. In: *Personality and social psychology bulletin* 44.3 (2018), pp. 332–344.
- [12] Timothy W Bickmore, Daniel Schulman, and Candace Sidner. “Automated interventions for multiple health behaviors using conversational agents”. In: *Patient education and counseling* 92.2 (2013), pp. 142–148.
- [13] Rúni Bláfoss et al. “Is fatigue after work a barrier for leisure-time physical activity? Cross-sectional study among 10,000 adults from the general working population”. In: *Scandinavian journal of public health* 47.3 (2019), pp. 383–391.
- [14] Mike GT Brannan et al. “Active 10—A new approach to increase physical activity in inactive people in England”. In: *Progress in Cardiovascular Diseases* 62.2 (2019), pp. 135–139.
- [15] Virginia Braun and Victoria Clarke. “Using thematic analysis in psychology”. In: *Qualitative research in psychology* 3.2 (2006), pp. 77–101.
- [16] Robert L Brennan and Dale J Prediger. “Coefficient kappa: Some uses, misuses, and alternatives”. In: *Educational and psychological measurement* 41.3 (1981), pp. 687–699.
- [17] Philip S Brenner and John D DeLamater. “Social desirability bias in self-reports of physical activity: is an exercise identity the culprit?” In: *Social Indicators Research* 117 (2014), pp. 489–504.
- [18] Roger Buehler and Dale Griffin. “Planning, personality, and prediction: The role of future focus in optimistic time predictions”. In: *Organizational Behavior and Human Decision Processes* 92.1-2 (2003), pp. 80–90.

- [19] Roger Buehler, Dale Griffin, and Michael Ross. "Exploring the" planning fallacy": Why people underestimate their task completion times." In: *Journal of personality and social psychology* 67.3 (1994), p. 366.
- [20] Fiona C Bull et al. "World Health Organization 2020 guidelines on physical activity and sedentary behaviour". In: *British Journal of Sports Medicine* 54.24 (2020), pp. 1451–1462. ISSN: 0306-3674. DOI: 10.1136/bjsports-2020-102955. eprint: <https://bjsm.bmj.com/content/54/24/1451.full.pdf>. URL: <https://bjsm.bmj.com/content/54/24/1451>.
- [21] Silke Burkert et al. "The interplay of dyadic and individual planning of pelvic-floor exercise in prostate-cancer patients following radical prostatectomy". In: *Journal of Behavioral Medicine* 35 (2012), pp. 305–317.
- [22] Meghan L Butryn et al. "Measuring the ability to tolerate activity-related discomfort: initial validation of the physical activity acceptance questionnaire (PAAQ)". In: *Journal of Physical Activity and Health* 12.5 (2015), pp. 717–726.
- [23] Rafael A Calvo et al. "Natural language processing in mental health applications using non-clinical texts". In: *Natural Language Engineering* 23.5 (2017), pp. 649–685.
- [24] Natasha Carraro and Patrick Gaudreau. "Spontaneous and experimentally induced action planning and coping planning for physical activity: A meta-analysis". In: *Psychology of Sport and Exercise* 14.2 (2013), pp. 228–248.
- [25] David Chapman and Leslie Pack Kaelbling. "Input Generalization in Delayed Reinforcement Learning: An Algorithm and Performance Comparisons." In: *Ijcai*. Vol. 91. 1991, pp. 726–731.
- [26] Richard A Chechile. *Bayesian statistics for experimental scientists: a general introduction using distribution-free methods*. MIT Press, 2020.
- [27] J Cohen. "Statistical power analysis for the behavioural sciences, 2nd edn New York". In: *NY: Lawrence Erlbaum Associates*. (1988).
- [28] Jacob Cohen. "A coefficient of agreement for nominal scales". In: *Educational and psychological measurement* 20.1 (1960), pp. 37–46.
- [29] Jacob Cohen. "A power primer." In: (2016).
- [30] Mark Conner, Tracy Sandberg, and Paul Norman. "Using action planning to promote exercise behavior". In: *Annals of Behavioral Medicine* 40.1 (2010), pp. 65–76.
- [31] Fred D Davis. "User acceptance of information systems: the technology acceptance model (TAM)". In: (1987).
- [32] Martin Dierikx, Nele Albers, and Willem-Paul Brinkman. *Daily Collaborative Personalized Step Goal-Setting with a Virtual Coach*. June 2023. DOI: 10.17605/OSF.IO/6JQPK. URL: osf.io/6jqpk.
- [33] Breno Quintella Farah et al. "Barriers to physical activity during the COVID-19 pandemic in adults: a cross-sectional study". In: *Sport Sciences for Health* 17 (2021), pp. 441–447.
- [34] Jasper Feine et al. "Gender bias in chatbot design". In: *Chatbot Research and Design: Third International Workshop, CONVERSATIONS 2019, Amsterdam, The Netherlands, November 19–20, 2019, Revised Selected Papers 3*. Springer. 2020, pp. 79–93.
- [35] Michela Ferron and Paolo Massa. "Transtheoretical model for designing technologies supporting an active lifestyle". In: *Proceedings of the Biannual Conference of the Italian Chapter of SIGCHI*. 2013, pp. 1–8.
- [36] Martin Fishbein and Icek Ajzen. "Belief, attitude, intention, and behavior: An introduction to theory and research". In: (1977).
- [37] Siska Fitrianie et al. *Artificial Social Agent Questionnaire Instrument*. en. 2023. DOI: 10.4121/19650846.v6. URL: <https://data.4tu.nl/datasets/d8e2c534-d192-4411-a1e2-5aeac97f0165/6>.
- [38] Siska Fitrianie et al. *Data and analysis underlying the research into the Artificial-Social-Agent Questionnaire: Establishing the long and short questionnaire versions*. en. 2023. DOI: 10.4121/19758436.v5. URL: <https://data.4tu.nl/datasets/fd344d58-1381-447c-a167-c6e53eaa0774/5>.

- [39] Siska Fitrianie et al. "The artificial-social-agent questionnaire: establishing the long and short questionnaire versions". In: *Proceedings of the 22nd ACM International Conference on Intelligent Virtual Agents*. 2022, pp. 1–8.
- [40] Jillian Francis et al. "Constructing questionnaires based on the theory of planned behaviour: A manual for health services researchers". In: (2004).
- [41] Tommy Gärling and Satoshi Fujii. "Structural equation modeling of determinants of planning". In: *Scandinavian Journal of Psychology* 43.1 (2002), pp. 1–8.
- [42] Emily Geisen. *Tips for measuring behavioral frequency*. URL: <https://www.qualtrics.com/blog/measuring-behavioral-frequency/> (visited on 06/14/2023).
- [43] Elizabeth Gerber et al. "Prototyping: facing uncertainty through small wins". In: *DS 58-9: Proceedings of ICED 09, the 17th International Conference on Engineering Design, Vol. 9, Human Behavior in Design, Palo Alto, CA, USA, 24.-27.08. 2009*. 2009, pp. 333–342.
- [44] Franklin N Glozah. "Exploring Ghanaian adolescents' meaning of health and wellbeing: A psychosocial perspective". In: *International Journal of Qualitative Studies on Health and Well-being* 10.1 (2015), p. 26370.
- [45] Gaston Godin. "The Godin-Shephard leisure-time physical activity questionnaire". In: *The Health & Fitness Journal of Canada* 4.1 (2011), pp. 18–22.
- [46] Martin S Hagger and Aleksandra Luszczynska. "Implementation intention and action planning interventions in health contexts: State of the research and proposals for the way forward". In: *Applied Psychology: Health and Well-Being* 6.1 (2014), pp. 1–47.
- [47] Mohand Hamadouche et al. "Comparison of value iteration, policy iteration and Q-Learning for solving decision-making problems". In: *2021 International Conference on Unmanned Aircraft Systems (ICUAS)*. IEEE. 2021, pp. 101–110.
- [48] Tim Harries et al. "Effectiveness of a smartphone app in increasing physical activity amongst male adults: a randomised controlled trial". In: *BMC public health* 16.1 (2016), pp. 1–10.
- [49] Takuya Hiraoka et al. "Reinforcement learning of cooperative persuasive dialogue policies using framing". In: *Proceedings of COLING 2014, the 25th International Conference on Computational Linguistics: Technical Papers*. 2014, pp. 1706–1717.
- [50] Beyza Hizli, Nele Albers, and Willem-Paul Brinkman. *Data and code underlying the master thesis: Goal-setting dialogue for physical activity with a virtual coach*. 2022. DOI: 10.4121/20047328.v1. URL: https://data.4tu.nl/articles/_/20047328/1.
- [51] Carl I Hovland and Walter Weiss. "The influence of source credibility on communication effectiveness". In: *Public opinion quarterly* 15.4 (1951), pp. 635–650.
- [52] Jan Keller et al. "Which characteristics of planning matter? Individual and dyadic physical activity plans and their effects on plan enactment". In: *Social Science & Medicine* 189 (2017), pp. 53–62.
- [53] Ewa Kulis et al. "Collaborative, dyadic, and individual planning and physical activity: A dyadic randomized controlled trial." In: *Health Psychology* 41.2 (2022), p. 134.
- [54] G Tarcan Kumkale and Dolores Albarracín. "The sleeper effect in persuasion: a meta-analytic review." In: *Psychological bulletin* 130.1 (2004), p. 143.
- [55] Margie E Lachman et al. "When adults don't exercise: Behavioral strategies to increase physical activity in sedentary middle-aged and older adults". In: *Innovation in aging* 2.1 (2018), igy007.
- [56] J Richard Landis and Gary G Koch. "The measurement of observer agreement for categorical data". In: *biometrics* (1977), pp. 159–174.
- [57] Charlotte Larmuseau, Piet Desmet, and Fien Depaepe. "Perceptions of instructional quality: Impact on acceptance and use of an online learning environment". In: *Interactive Learning Environments* 27.7 (2019), pp. 953–964.
- [58] Gary P Latham and Edwin A Locke. "Self-regulation through goal setting". In: *Organizational behavior and human decision processes* 50.2 (1991), pp. 212–247.

- [59] Mackenzie Leake et al. "Generating Audio-Visual Slideshows from Text Articles Using Word Concreteness." In: *CHI*. Vol. 20. 2020, pp. 25–30.
- [60] I-Min Lee and David M Buchner. "The importance of walking to public health". In: *Medicine & Science in Sports & Exercise* 40.7 (2008), S512–S518.
- [61] Stephanie Anna Lenzen et al. "Disentangling self-management goal setting and action planning: A scoping review". In: *PLoS One* 12.11 (2017), e0188822.
- [62] Edwin A Locke and Gary P Latham. "New directions in goal-setting theory". In: *Current directions in psychological science* 15.5 (2006), pp. 265–268.
- [63] Edwin A Locke, Gary P Latham, and Miriam Erez. "The determinants of goal commitment". In: *Academy of management review* 13.1 (1988), pp. 23–39.
- [64] Edwin A. Locke. "Motivation through conscious goal setting". In: *Applied and Preventive Psychology* 5.2 (1996), pp. 117–124. ISSN: 0962-1849. DOI: [https://doi.org/10.1016/S0962-1849\(96\)80005-9](https://doi.org/10.1016/S0962-1849(96)80005-9). URL: <https://www.sciencedirect.com/science/article/pii/S0962184996800059>.
- [65] Jemma Looyestyn et al. "Does gamification increase engagement with online programs? A systematic review". In: *PloS one* 12.3 (2017), e0173403.
- [66] Kate Lorig et al. "The components of action planning and their associations with behavior and health outcomes". In: *Chronic Illness* 10.1 (2014), pp. 50–59.
- [67] Tiffany Christina Luo et al. "Promoting physical activity through conversational agents: mixed methods systematic review". In: *Journal of Medical Internet Research* 23.9 (2021), e25486.
- [68] Carol Ann Maher et al. "A physical activity and diet program delivered by artificially intelligent virtual health coach: proof-of-concept study". In: *JMIR mHealth and uHealth* 8.7 (2020), e17558.
- [69] Anna V. Mattioli et al. "Quarantine during COVID-19 outbreak: Changes in diet and physical activity increase the risk of cardiovascular disease". In: *Nutrition, Metabolism and Cardiovascular Diseases* 30.9 (2020), pp. 1409–1417. ISSN: 0939-4753. DOI: <https://doi.org/10.1016/j.numecd.2020.05.020>. URL: <https://www.sciencedirect.com/science/article/pii/S0939475320302131>.
- [70] A Peter McGraw, Barbara A Mellers, and Ilana Ritov. "The affective costs of overconfidence". In: *Journal of Behavioral Decision Making* 17.4 (2004), pp. 281–295.
- [71] Anouk Middelweerd et al. "Apps to promote physical activity among adults: a review and content analysis". In: *International journal of behavioral nutrition and physical activity* 11.1 (2014), pp. 1–9.
- [72] Shiwali Mohan, Anusha Venkatakrishnan, and Andrea L Hartzler. "Designing an AI health coach and studying its utility in promoting regular aerobic exercise". In: *ACM Transactions on Interactive Intelligent Systems (TiiS)* 10.2 (2020), pp. 1–30.
- [73] Suzanne Morony et al. "Predicting achievement: Confidence vs self-efficacy, anxiety, and self-concept in Confucian and European countries". In: *International Journal of Educational Research* 58 (2013), pp. 79–96.
- [74] Ryo Nagata, Tomoya Hashiguchi, and Driss Sadoun. "Is the Simplest Chatbot Effective in English Writing Learning Assistance?" In: *Computational Linguistics: 16th International Conference of the Pacific Association for Computational Linguistics, PACLING 2019, Hanoi, Vietnam, October 11–13, 2019, Revised Selected Papers 16*. Springer. 2020, pp. 245–256.
- [75] Michael Neff et al. "Evaluating the effect of gesture and language on personality perception in conversational agents". In: *Intelligent Virtual Agents: 10th International Conference, IVA 2010, Philadelphia, PA, USA, September 20–22, 2010. Proceedings 10*. Springer. 2010, pp. 222–235.
- [76] Huy Anh Nguyen, Jake M Hofman, and Daniel G Goldstein. "Round numbers can sharpen cognition". In: *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 2022, pp. 1–15.
- [77] David Nichols. *Coloring for Colorblindness*. URL: <https://davidmathlogic.com/colorblind/> (visited on 05/30/2023).

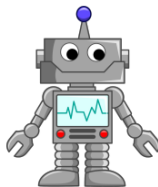
- [78] Claudio Nigg. *Exercise: Stages of Change (Short Form)*. URL: <https://web.uri.edu/cprc/measures/exercise/stages-of-change-short-form/> (visited on 06/14/2023).
- [79] Travis E Oliphant. "A Bayesian perspective on estimating mean, variance, and standard-deviation from data". In: (2006).
- [80] Richard E Petty et al. *The elaboration likelihood model of persuasion*. Springer, 1986.
- [81] Theda Radtke et al. "A cluster randomized controlled trial comparing the effectiveness of an individual planning intervention with collaborative planning in adolescent friendship dyads to enhance physical activity (TWOgether)". In: *BMC public health* 18.1 (2018), pp. 1–11.
- [82] Stephanie A Robinson et al. "Time for change: using implementation intentions to promote physical activity in a randomised pilot trial". In: *Psychology & health* 34.2 (2019), pp. 232–254.
- [83] Amelia Romeo et al. "Can smartphone apps increase physical activity? Systematic review and meta-analysis". In: *Journal of medical Internet research* 21.3 (2019), e12053.
- [84] James A Russell. "A circumplex model of affect." In: *Journal of personality and social psychology* 39.6 (1980), p. 1161.
- [85] J Edward Russo, Paul JH Schoemaker, et al. "Managing overconfidence". In: *Sloan management review* 33.2 (1992), pp. 7–17.
- [86] Lesley Scobbie, Diane Dixon, and Sally Wyke. "Goal setting and action planning in the rehabilitation setting: development of a theoretically informed practice framework". In: *Clinical Rehabilitation* 25.5 (2011), pp. 468–482.
- [87] Eunjin Seo et al. "The effects of goal origin and implementation intentions on goal commitment, effort, and performance". In: *The Journal of Experimental Education* 86.3 (2018), pp. 386–401.
- [88] Nancy E Sherwood and Robert W Jeffery. "The behavioral determinants of exercise: implications for physical activity interventions". In: *Annual review of nutrition* 20.1 (2000), pp. 21–44.
- [89] Falko F Sniehotta, Urte Scholz, and Ralf Schwarzer. "Action plans and coping plans for physical exercise: A longitudinal intervention study in cardiac rehabilitation". In: *British journal of health psychology* 11.1 (2006), pp. 23–37.
- [90] Falko F Sniehotta et al. "Action planning and coping planning for long-term lifestyle change: theory and assessment". In: *European Journal of Social Psychology* 35.4 (2005), pp. 565–576.
- [91] Andrei Stefan. *PerfectFit-project/rasa_planning_agent: version1*. Version version1. Nov. 2023. DOI: 10.5281/zenodo.10126244. URL: <https://doi.org/10.5281/zenodo.10126244>.
- [92] Andrei Stefan, Nele Albers, and Willem-Paul Brinkman. *Data and analysis code for the Masters Thesis titled "Dyadic Physical Activity Planning with a Virtual Coach: Using Reinforcement Learning to Select Persuasive Strategies"*. 2023. DOI: 10.4121/2796f502-0610-4a7d-a8ee-ebc36639e0b1.
- [93] Andrei Stefan, Nele Albers, and Willem-Paul Brinkman. *Setting Goals For Physical Activity Through Dialogue*. July 2023. DOI: 10.17605/OSF.IO/8ADP9. URL: osf.io/nqawd.
- [94] Martin Stevens et al. "The Groningen Enjoyment Questionnaire: a measure of enjoyment in leisure-time physical activity". In: *Perceptual and motor skills* 90.2 (2000), pp. 601–604.
- [95] Stephanie Stockwell et al. "Changes in physical activity and sedentary behaviours from before to during the COVID-19 pandemic lockdown: a systematic review". In: *BMJ Open Sport & Exercise Medicine* 7.1 (2021). DOI: 10.1136/bmjsem-2020-000960. eprint: <https://bmjopensem.bmj.com/content/7/1/e000960.full.pdf>. URL: <https://bmjopensem.bmj.com/content/7/1/e000960>.
- [96] Ryan J Summitt et al. "Shoulder injuries in individuals who participate in CrossFit training". In: *Sports health* 8.6 (2016), pp. 541–546.
- [97] Shelley E Taylor et al. "Harnessing the imagination: Mental simulation, self-regulation, and coping." In: *American psychologist* 53.4 (1998), p. 429.
- [98] Xintao Tian et al. "Considering interaction sequence of historical items for conversational recommender system". In: *Database Systems for Advanced Applications: 26th International Conference, DASFAA 2021, Taipei, Taiwan, April 11–14, 2021, Proceedings, Part III* 26. Springer. 2021, pp. 115–131.

- [99] Nhat Tran, Malihe Alikhani, and Diane Litman. “How to ask for donations? learning user-specific persuasive dialogue policies through online interactions”. In: *Proceedings of the 30th ACM Conference on User Modeling, Adaptation and Personalization*. 2022, pp. 12–22.
- [100] Catrine Tudor-Locke, William D Johnson, and Peter T Katzmarzyk. “Accelerometer-determined steps per day in US adults.” In: *Medicine and science in sports and exercise* 41.7 (2009), pp. 1384–1391.
- [101] Catrine Tudor-Locke et al. “Walking cadence (steps/min) and intensity in 21–40 year olds: CADENCE-adults”. In: *International Journal of Behavioral Nutrition and Physical Activity* 16.1 (2019), pp. 1–11.
- [102] VENNGAGE. *Accessible color palette generator*. URL: <https://venngage.com/tools/accessible-color-palette-generator> (visited on 05/30/2023).
- [103] Qing Wang et al. “Diet and physical activity apps: perceived effectiveness by app users”. In: *JMIR mHealth and uHealth* 4.2 (2016), e5114.
- [104] Darren ER Warburton, Crystal Whitney Nicol, and Shannon SD Bredin. “Health benefits of physical activity: the evidence”. In: *Cmaj* 174.6 (2006), pp. 801–809.
- [105] Apichai Wattanapisit and Sanhapan Thanamee. “Evidence behind 10,000 steps walking”. In: *Journal of Health Research* 31.3 (2017), pp. 241–248.
- [106] Klaus Weber et al. “Predicting persuasive effectiveness for multimodal behavior adaptation using bipolar weighted argument graphs”. In: (2020).
- [107] Stefanie L Williams and David P French. “What are the most effective intervention techniques for changing physical activity self-efficacy and physical activity behaviour—and are they the same?” In: *Health education research* 26.2 (2011), pp. 308–322.
- [108] Kefan Xu, Xinghui Yan, and Mark W Newman. “Understanding People’s Experience for Physical Activity Planning and Exploring the Impact of Historical Records on Plan Creation and Execution”. In: *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 2022, pp. 1–15.
- [109] Jochen P Ziegelmann, Sonia Lippke, and Ralf Schwarzer. “Adoption and maintenance of physical activity: Planning interventions in young, middle-aged, and older adults”. In: *Psychology & Health* 21.2 (2006), pp. 145–163.

A

Scenarios used in the expert discussions

Scenario 1: amount of physical activity (per day) and breaks between days



Moving on to creating the plan. Considering you want to do 100 minutes of running per week, I suggest the following plan:

Note: we assume the user doesn't have a preference for using as few days as possible

Option A: More activity, but less days and more breaks

Option B: Less activity, but more days and less breaks

	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
Morning (6:00 - 11:00)							
Midday (11:00-14:00)							
Afternoon (14:00-18:00)							
Evening (18:00-22:00)			Run 33 min		Run 33 min	Run 33 min	

	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
Morning (6:00 - 11:00)							
Midday (11:00-14:00)							
Afternoon (14:00-18:00)							
Evening (18:00-22:00)	Run 20 min	Run 20 min	Run 20 min			Run 20 min	Run 20 min

Figure A.1: The first scenario used in the expert discussions

Scenario 2: Breaks during the same day

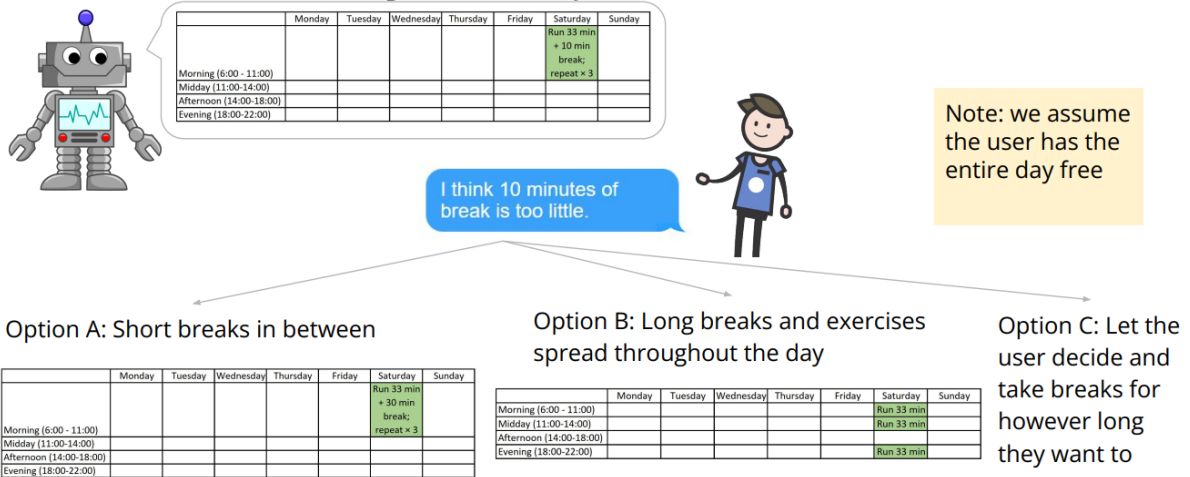


Figure A.2: The second scenario used in the expert discussions

Scenario 3: dealing with constraints (such as daily limit or breaks)

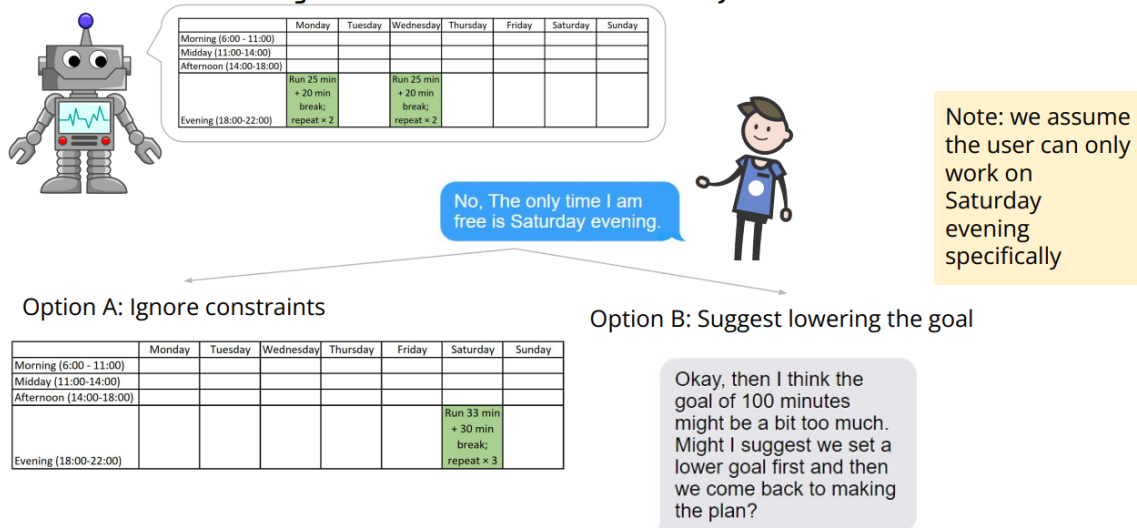
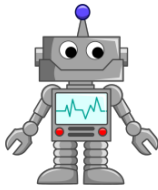


Figure A.3: The third scenario used in the expert discussions

Scenario 4: amount of detail



Moving on to creating the plan. Considering you want to do 100 minutes of running per week, I suggest the following plan:

Option C: also place

	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
Morning (6:00 - 11:00)							
Midday (11:00-14:00)							
Afternoon (14:00-18:00)							
Evening (18:00-22:00)			Run 33 min from 20:00 at the park		Run 33 min from 19:00 at the park	Run 33 min from 20:30 at the gym	

Option A: vague, leaving the details up to the user

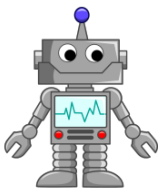
	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
Morning (6:00 - 11:00)							
Midday (11:00-14:00)							
Afternoon (14:00-18:00)							
Evening (18:00-22:00)			Run 33 min		Run 33 min	Run 33 min	

Option B: also time

	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
Morning (6:00 - 11:00)							
Midday (11:00-14:00)							
Afternoon (14:00-18:00)							
Evening (18:00-22:00)			Run 33 min from 20:00		Run 33 min from 19:00	Run 33 min from 20:30	

Figure A.4: The fourth scenario used in the expert discussions

Scenario 5: amount of detail over time



Moving on to creating the plan. Considering you want to do 100 minutes of running per week, I suggest the following plan:

Option B: detailed to the week/month level

Option A: detailed to the day level

Week 1	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
Morning (6:00 - 11:00)							
Midday (11:00-14:00)							
Afternoon (14:00-18:00)							
Evening (18:00-22:00)			Run 33 min		Run 33 min	Run 33 min	
Week 2	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
Morning (6:00 - 11:00)							
Midday (11:00-14:00)							
Afternoon (14:00-18:00)							
Evening (18:00-22:00)			Run 36 min		Run 36 min	Run 36 min	
Week 3	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
Morning (6:00 - 11:00)							
Midday (11:00-14:00)							
Afternoon (14:00-18:00)							
Evening (18:00-22:00)			Run 40 min		Run 40 min	Run 40 min	

Week 1	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
Morning (6:00 - 11:00)							
Midday (11:00-14:00)							
Afternoon (14:00-18:00)							
Evening (18:00-22:00)			Run 33 min		Run 33 min	Run 33 min	
Week 2	Keep goal of 100 minutes per week						
Week 3	Increase goal to 110 minutes per week						
Week 4	Increase goal to 115 minutes per week						
Month 2	Increase goal to 125 minutes per week						
Month 3	Increase goal to 150 minutes per week						

Figure A.5: The fifth scenario used in the expert discussions

Scenario 6: timing conflicts

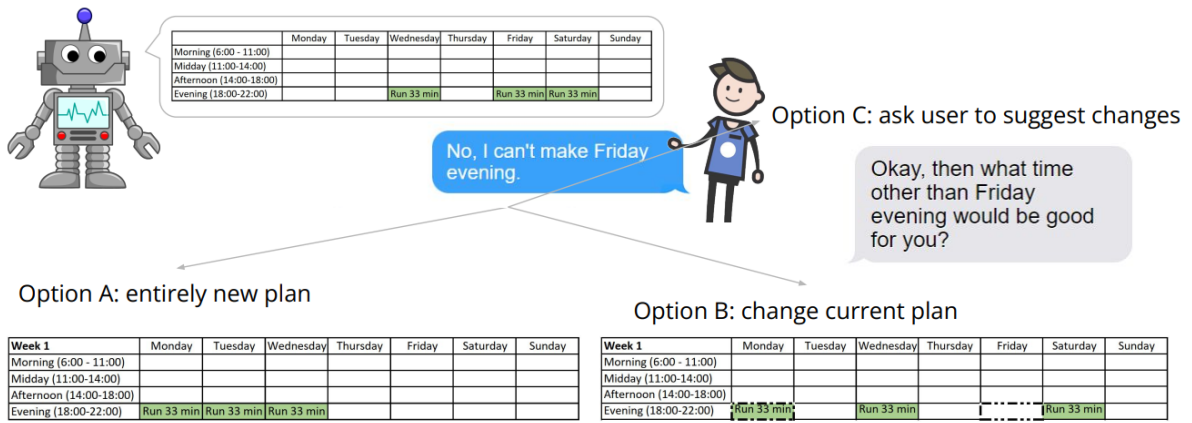


Figure A.6: The sixth scenario used in the expert discussions

Scenario 7: presentation

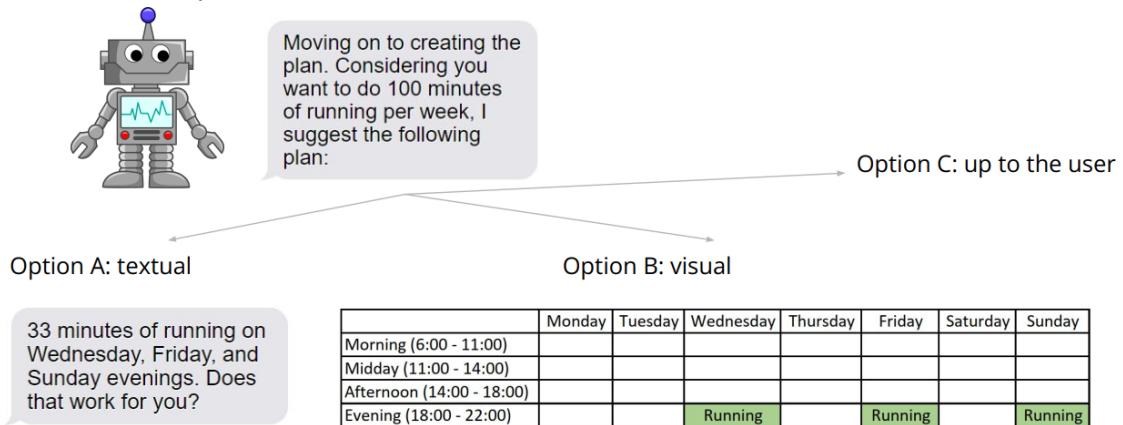
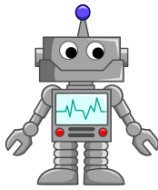


Figure A.7: The seventh scenario used in the expert discussions

Scenario 8: time frame for planning



Moving on to creating the plan. Considering you want to do 100 minutes of running per week, I suggest the following plan:

Option A: one week

	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
Morning (6:00 - 11:00)							
Midday (11:00-14:00)							
Afternoon (14:00-18:00)							
Evening (18:00-22:00)			Run 33 min		Run 33 min	Run 33 min	

Option B: many weeks

Week 1	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
Morning (6:00 - 11:00)							
Midday (11:00-14:00)							
Afternoon (14:00-18:00)							
Evening (18:00-22:00)			Run 33 min		Run 33 min	Run 33 min	
Week 2	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
Morning (6:00 - 11:00)	Run 33 min			Run 33 min			Run 33 min
Midday (11:00-14:00)							
Afternoon (14:00-18:00)							
Evening (18:00-22:00)							
Week 3	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
Morning (6:00 - 11:00)		Run 37 min					
Midday (11:00-14:00)							
Afternoon (14:00-18:00)				Run 37 min		Run 37 min	
Evening (18:00-22:00)							
Week 4	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
Morning (6:00 - 11:00)							
Midday (11:00-14:00)							
Afternoon (14:00-18:00)							
Evening (18:00-22:00)			Run 25 min + 20 min break; repeat x 2				Run 30 min + 30 min break; repeat x 2
Week 5	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
Morning (6:00 - 11:00)				Run 30 min			
Midday (11:00-14:00)							
Afternoon (14:00-18:00)							
Evening (18:00-22:00)	Run 30 min	Run 30 min				Run 30 min	

Figure A.8: The eighth scenario used in the expert discussions

B

Implementation of the persuasive strategies

B.1. Proposing to make changes to the plan

When this persuasive strategy is used, the virtual coach informs the person that they have the opportunity to change the times when they have to go for walks. They can do this by clicking the time slots displayed in the table on the left side of their screen. This can be seen in Figure B.1

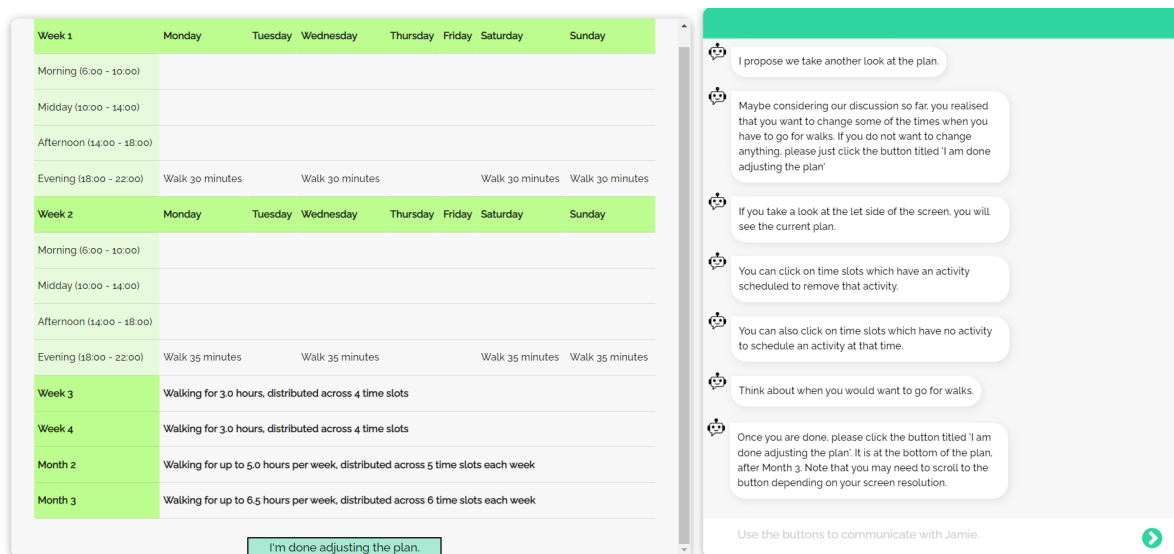


Figure B.1: The chat for the action of proposing to make changes to the plan

B.2. Explain why planning is useful

The virtual coach presents the person with a few reasons why planning can help people do physical activity. The reasons are the following:

- It can help keep me motivated to go for walks,
- It can help me create a habit of taking walks regularly,
- It can help me visualise the road towards the goal,
- It can help me deal with obstacles.

People are asked to pick the one they think would be the most beneficial for them and to type in their answer (Figure B.2), after which the virtual coach shows the corresponding explanation (Figure B.3).

The reasons and the corresponding explanations shown to people are based on research by Sniehotta et al. [90] and Sniehotta, Scholz, and Schwarzer [89], and on our discussions with an expert in health psychology.

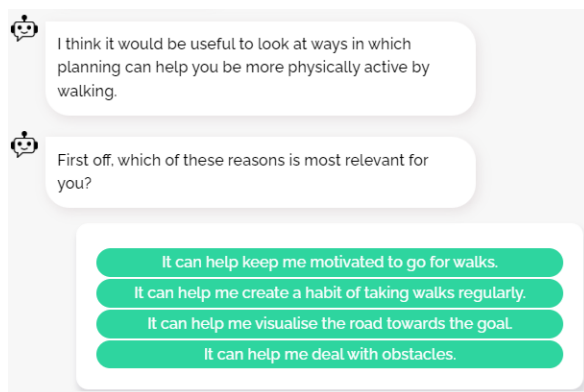


Figure B.2: The first part of the chat for the action of explaining why planning is useful

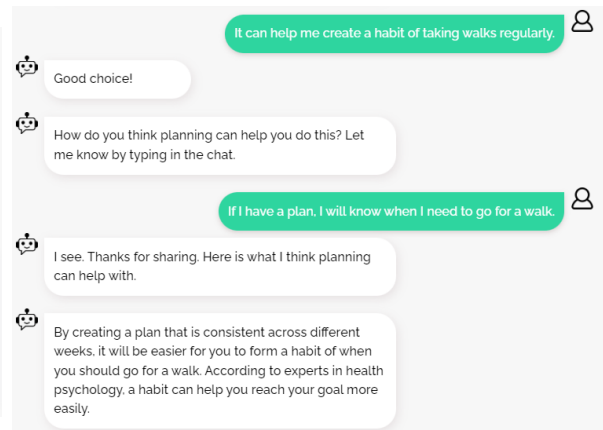


Figure B.3: The second part of the chat for the action of explaining why planning is useful

B.3. Identify barriers and Deal with barriers

The virtual coach first helps people identify barriers, by providing a list of common barriers for planning for physical activity:

- Lack of time,
- Lack of energy,
- Disliking being in the presence of others when exercising,
- Lack of equipment,
- Family obligations.

When identifying barriers, the virtual coach presents the common options shown above, along with the option "I have a different barrier". The person is asked to pick the barrier type they think could hinder them the most (Figure B.4), and then type in the chat their barrier in detail (Figure B.5).

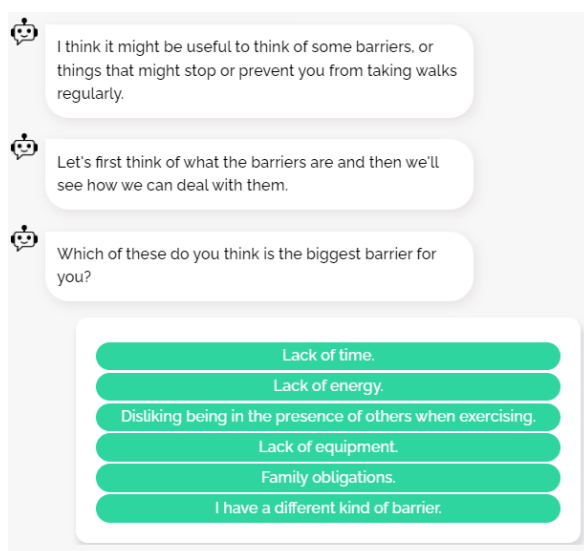


Figure B.4: The first part of the chat for the action of identifying barriers

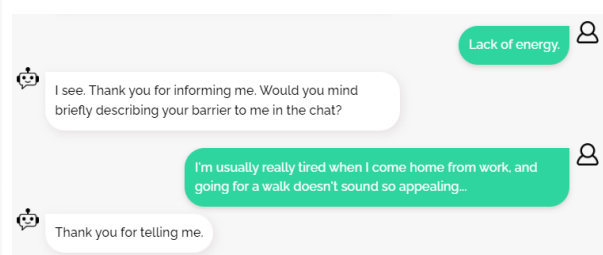


Figure B.5: The second part of the chat for the action of identifying barriers

When dealing with barriers, the virtual coach first informs the person that it can provide suggestions, but that it is up to the person to find a good way of dealing with their barrier (Figure B.6). Then, the virtual coach repeats what the person typed in previously as their barrier and asks them to type their approach for dealing with their barrier (Figure B.7). After this, if the barrier was not "I have a different barrier", a way of dealing with the specific barrier type is presented, and the virtual coach asks the person to type their strategy after having read its suggestion, since it might be different from what it was initially (Figure B.8).

The first three barriers and the methods for dealing with them are selected based on research by Baillot et al. [7], Bláfoss et al. [13], and Glozah [44], while the fourth barrier and the method of dealing with it is additionally based on research by Farah et al. [33], and the fifth barrier and the way of dealing with it is based on research by Bantham et al. [9].

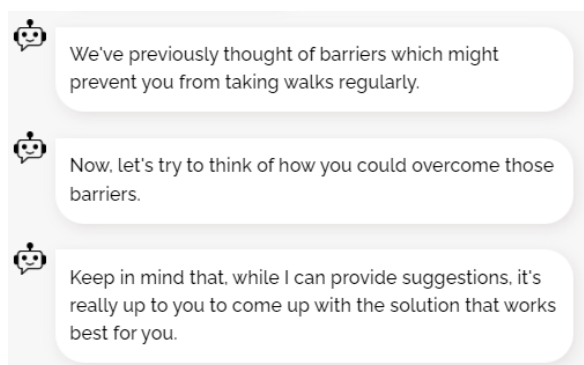


Figure B.6: The first part of the chat for the action of dealing with barriers

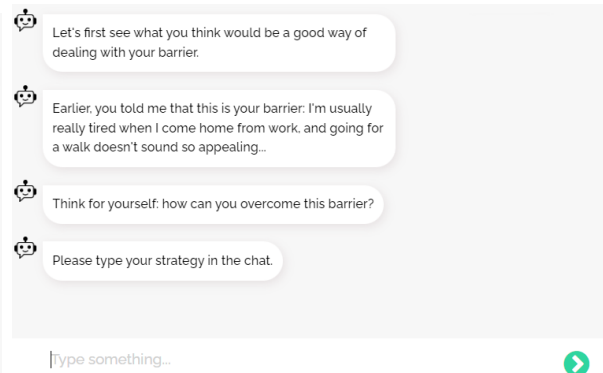


Figure B.7: The second part of the chat for the action of dealing with barriers

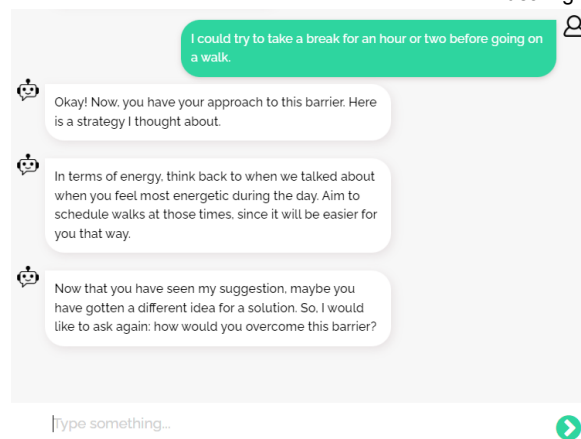


Figure B.8: The third part of the chat for the action of dealing with barriers

B.4. Show testimonials from other people who followed plans for physical activity

Two random testimonials are shown when this action is done. The person is asked to read each testimonial and then type in the chat something they could take away from it, for every testimonial (Figure B.9).

The testimonials shown are taken from the research by Albers et al. [5], which has a publicly available dataset [50]. The 72 testimonials were coded by two people with a background in computer science, in 4 categories (or combinations of categories): testimonials which are not related to planning, testimonials which relate to people creating plans for physical activity, testimonials which relate to people following plans for physical activity, testimonials which relate to people using plans for physical activity. The testimonials where both coders agreed on the categories (or combination of categories)

were selected and, of those testimonials, those which feature people following plans for physical activity were the ones picked to be shown to people. Thus, 8 testimonials were selected and two of them were randomly chosen every time the virtual coach used this persuasive strategy.

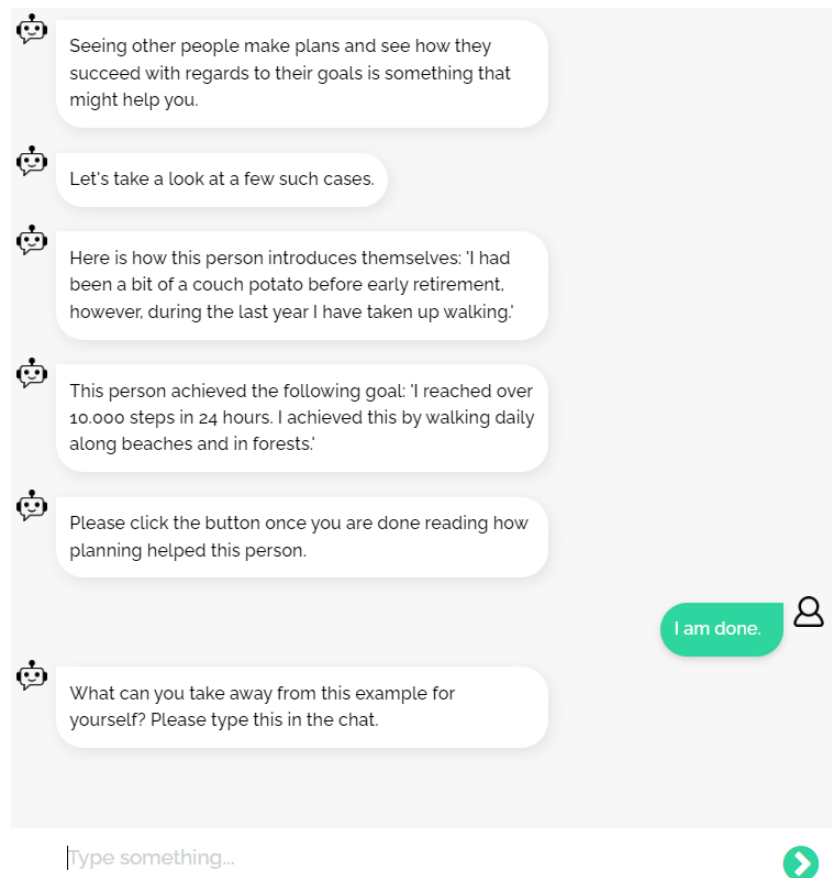


Figure B.9: The chat for the action of showing testimonials

C

Distribution of the data

Table C.1: Distribution of the actions taken per reduced state

State	Action					Total
	changes to plan	explain planning	identify barriers	deal with barriers	show testimonials	
['0', '0']	51	25	22	23	27	148
['0', '1']	18	8	13	11	11	61
['1', '0']	23	14	17	10	11	75
['1', '1']	81	39	33	41	41	235
Total	173	86	85	85	90	519

D

Detailed description of the process of selecting features and values using the G-algorithm

The purpose of the G-algorithm was to provide a basis for selecting only two of the three possible features which were part of the person's situation (confidence, perceived usefulness, and attitude).

Furthermore, rather than only checking if including a feature was relevant, the algorithm checked if including the feature with each of the possible number of values was relevant. In other words, for each feature, the algorithm checked if including the feature with two, three, or four values split the data such that the resulting subsets were significantly different from each other. Determining how many values each feature should have and how to make this split was necessary, since the data for the states was not distributed uniformly, as seen in the pilot study.

Table D.1: Results of running the G algorithm for predicting Q-values.

Iteration number	Feature	Number of values	<i>p</i> -value
1	confidence	2	0.0066
		3	0.0051
		4	0.0004
1	perceived usefulness	2	0.0080
		3	0.0160
		4	0.0648
1	attitude	2	0.0522
		3	0.0815
		4	0.2467
Feature selected: confidence, with two values.			
2	perceived usefulness	2	< 0.0001
		3	0.0011
		4	0.0064
2	attitude	2	0.0017
		3	0.0127
		4	0.0086
Feature selected: perceived usefulness, with two values.			

In table, D.1, we can see that confidence with two values is selected in the first iteration, since it has the smallest *p*-value. Then, in the second iteration, perceived usefulness with two values has the smallest *p*-value, so it is picked as the second feature.

E

Optimal policy

Table E.1: Optimal policy for the model using the first reward function. States are of the form [number of changes made to the plan, 'confidence', 'perceived usefulness', explaining why planning is useful was done, identifying barriers was done, dealing with barriers was done, showing testimonials from other people was done]. For all other possible states not listed in the table, the optimal action is to pick a random action out of the possible actions for the specific state.

State	Action
[0, '0', '0', True, False, False, False]	changes to plan
[0, '0', '1', True, False, False, False]	changes to plan
[0, '1', '0', True, False, False, False]	changes to plan
[0, '1', '1', True, False, False, False]	changes to plan
[0, '0', '0', False, True, False, False]	changes to plan
[0, '0', '1', False, True, False, False]	changes to plan
[0, '1', '0', False, True, False, False]	changes to plan
[0, '1', '1', False, True, False, False]	changes to plan
[0, '0', '0', False, False, False, True]	changes to plan
[0, '0', '1', False, False, False, True]	changes to plan
[0, '1', '0', False, False, False, True]	changes to plan
[0, '1', '1', False, False, False, True]	changes to plan
[0, '0', '0', False, False, False, False]	identify barriers
[0, '0', '1', False, False, False, False]	explain planning
[0, '1', '0', False, False, False, False]	explain planning
[0, '1', '1', False, False, False, False]	explain planning
[1, '0', '0', True, True, True, True]	changes to plan
[1, '0', '1', True, True, True, True]	changes to plan
[1, '1', '0', True, True, True, True]	changes to plan
[1, '1', '1', True, True, True, True]	changes to plan
[1, '0', '1', True, True, True, False]	changes to plan
[1, '1', '0', True, True, True, False]	changes to plan
[1, '0', '1', True, True, False, True]	changes to plan
[1, '1', '0', True, True, False, True]	changes to plan
[1, '0', '0', True, True, False, False]	deal with barriers
[1, '0', '1', True, True, False, False]	deal with barriers
[1, '1', '0', True, True, False, False]	deal with barriers
[1, '1', '1', True, True, False, False]	deal with barriers
[1, '0', '0', True, False, False, True]	identify barriers
[1, '0', '1', True, False, False, True]	identify barriers
[1, '1', '0', True, False, False, True]	identify barriers
[1, '1', '1', True, False, False, True]	identify barriers

[1, '0', '0', True, False, False, False]	identify barriers
[1, '0', '1', True, False, False, False]	show testimonials
[1, '1', '0', True, False, False, False]	show testimonials
[1, '1', '1', True, False, False, False]	show testimonials
[1, '0', '1', False, True, True, True]	changes to plan
[1, '1', '0', False, True, True, True]	changes to plan
[1, '0', '0', False, True, True, False]	explain planning
[1, '0', '1', False, True, True, False]	explain planning
[1, '1', '0', False, True, True, False]	show testimonials
[1, '1', '1', False, True, True, False]	explain planning
[1, '0', '0', False, True, False, True]	deal with barriers
[1, '0', '1', False, True, False, True]	explain planning
[1, '1', '0', False, True, False, True]	deal with barriers
[1, '1', '1', False, True, False, True]	deal with barriers
[1, '0', '0', False, True, False, False]	deal with barriers
[1, '0', '1', False, True, False, False]	deal with barriers
[1, '1', '0', False, True, False, False]	deal with barriers
[1, '1', '1', False, True, False, False]	deal with barriers
[1, '0', '0', False, False, False, True]	identify barriers
[1, '0', '1', False, False, False, True]	explain planning
[1, '1', '0', False, False, False, True]	explain planning
[1, '1', '1', False, False, False, True]	explain planning
[1, '0', '0', False, False, False, False]	identify barriers
[1, '0', '1', False, False, False, False]	explain planning
[1, '1', '0', False, False, False, False]	show testimonials
[1, '1', '1', False, False, False, False]	identify barriers
[2, '0', '0', True, True, True, False]	show testimonials
[2, '0', '1', True, True, True, False]	show testimonials
[2, '1', '0', True, True, True, False]	show testimonials
[2, '0', '0', True, True, False, True]	deal with barriers
[2, '0', '1', True, True, False, True]	deal with barriers
[2, '1', '0', True, True, False, True]	deal with barriers
[2, '0', '0', True, True, False, False]	deal with barriers
[2, '1', '0', True, True, False, False]	deal with barriers
[2, '0', '0', True, False, False, True]	identify barriers
[2, '0', '1', True, False, False, True]	identify barriers
[2, '1', '0', True, False, False, True]	identify barriers
[2, '0', '0', True, False, False, False]	identify barriers
[2, '0', '1', True, False, False, False]	show testimonials
[2, '1', '0', True, False, False, False]	show testimonials
[2, '0', '0', False, True, True, True]	explain planning
[2, '0', '1', False, True, True, True]	explain planning
[2, '1', '0', False, True, True, True]	explain planning
[2, '0', '0', False, True, True, False]	explain planning
[2, '0', '1', False, True, True, False]	explain planning
[2, '1', '0', False, True, True, False]	show testimonials
[2, '0', '0', False, True, False, True]	deal with barriers
[2, '0', '1', False, True, False, True]	explain planning
[2, '1', '0', False, True, False, True]	deal with barriers
[2, '0', '0', False, True, False, False]	deal with barriers
[2, '0', '1', False, True, False, False]	explain planning
[2, '1', '0', False, True, False, False]	deal with barriers
[2, '0', '0', False, False, False, True]	identify barriers

[2, '0', '1', False, False, False, True]	explain planning
[2, '1', '0', False, False, False, True]	explain planning

F

Worst policy

Table F.1: Worst policy for the model using the first reward function. States are of the form [number of changes made to the plan, 'confidence', 'perceived usefulness', explaining why planning is useful was done, identifying barriers was done, dealing with barriers was done, showing testimonials from other people was done]. For all other possible states not listed in the table, the optimal action is to pick a random action out of the possible actions for the specific state.

State	Action
[0, '0', '0', True, False, False, False]	changes to plan
[0, '0', '1', True, False, False, False]	changes to plan
[0, '1', '0', True, False, False, False]	changes to plan
[0, '1', '1', True, False, False, False]	changes to plan
[0, '0', '0', False, True, False, False]	changes to plan
[0, '0', '1', False, True, False, False]	changes to plan
[0, '1', '0', False, True, False, False]	changes to plan
[0, '1', '1', False, True, False, False]	changes to plan
[0, '0', '0', False, False, False, True]	changes to plan
[0, '0', '1', False, False, False, True]	changes to plan
[0, '1', '0', False, False, False, True]	changes to plan
[0, '1', '1', False, False, False, True]	changes to plan
[0, '0', '0', False, False, False, False]	changes to plan
[0, '0', '1', False, False, False, False]	changes to plan
[0, '1', '0', False, False, False, False]	changes to plan
[0, '1', '1', False, False, False, False]	changes to plan
[1, '0', '0', True, True, True, True]	changes to plan
[1, '0', '1', True, True, True, True]	changes to plan
[1, '1', '0', True, True, True, True]	changes to plan
[1, '1', '1', True, True, True, True]	changes to plan
[1, '0', '0', True, True, True, False]	changes to plan
[1, '0', '1', True, True, True, False]	changes to plan
[1, '1', '0', True, True, True, False]	changes to plan
[1, '1', '1', True, True, True, False]	changes to plan
[1, '0', '0', True, True, False, True]	changes to plan
[1, '0', '1', True, True, False, True]	changes to plan
[1, '1', '0', True, True, False, True]	changes to plan
[1, '1', '1', True, True, False, True]	changes to plan
[1, '0', '1', True, True, False, False]	changes to plan
[1, '1', '0', True, True, False, False]	changes to plan
[1, '0', '0', True, False, False, True]	changes to plan
[1, '0', '1', True, False, False, True]	changes to plan

[1, '1', '0', True, False, False, True]	changes to plan
[1, '1', '1', True, False, False, True]	changes to plan
[1, '0', '0', True, False, False, False]	show testimonials
[1, '0', '1', True, False, False, False]	identify barriers
[1, '1', '0', True, False, False, False]	identify barriers
[1, '1', '1', True, False, False, False]	identify barriers
[1, '0', '0', False, True, True, True]	changes to plan
[1, '0', '1', False, True, True, True]	changes to plan
[1, '1', '0', False, True, True, True]	changes to plan
[1, '1', '1', False, True, True, True]	changes to plan
[1, '0', '0', False, True, True, False]	show testimonials
[1, '0', '1', False, True, True, False]	changes to plan
[1, '1', '0', False, True, True, False]	changes to plan
[1, '1', '1', False, True, True, False]	changes to plan
[1, '0', '1', False, True, False, True]	changes to plan
[1, '1', '0', False, True, False, True]	changes to plan
[1, '0', '1', False, True, False, False]	show testimonials
[1, '1', '0', False, True, False, False]	explain planning
[1, '0', '0', False, False, False, True]	explain planning
[1, '0', '1', False, False, False, True]	identify barriers
[1, '1', '0', False, False, False, True]	identify barriers
[1, '1', '1', False, False, False, True]	identify barriers
[1, '0', '1', False, False, False, False]	identify barriers
[1, '1', '0', False, False, False, False]	identify barriers
[2, '0', '0', True, True, True, False]	show testimonials
[2, '0', '1', True, True, True, False]	show testimonials
[2, '1', '0', True, True, True, False]	show testimonials
[2, '1', '1', True, True, True, False]	show testimonials
[2, '0', '0', True, True, False, True]	deal with barriers
[2, '0', '1', True, True, False, True]	deal with barriers
[2, '1', '0', True, True, False, True]	deal with barriers
[2, '1', '1', True, True, False, True]	deal with barriers
[2, '0', '0', True, True, False, False]	show testimonials
[2, '1', '0', True, True, False, False]	show testimonials
[2, '0', '0', True, False, False, True]	identify barriers
[2, '0', '1', True, False, False, True]	identify barriers
[2, '1', '0', True, False, False, True]	identify barriers
[2, '1', '1', True, False, False, True]	identify barriers
[2, '0', '0', True, False, False, False]	show testimonials
[2, '0', '1', True, False, False, False]	identify barriers
[2, '1', '0', True, False, False, False]	identify barriers
[2, '0', '0', False, True, True, True]	explain planning
[2, '0', '1', False, True, True, True]	explain planning
[2, '1', '0', False, True, True, True]	explain planning
[2, '1', '1', False, True, True, True]	explain planning
[2, '0', '0', False, True, True, False]	show testimonials
[2, '0', '1', False, True, True, False]	show testimonials
[2, '1', '0', False, True, True, False]	explain planning
[2, '0', '0', False, True, False, True]	explain planning
[2, '0', '1', False, True, False, True]	deal with barriers
[2, '1', '0', False, True, False, True]	explain planning
[2, '1', '0', False, True, False, False]	explain planning
[2, '0', '0', False, False, False, True]	explain planning

[2, '0', '1', False, False, False, True]	identify barriers
[2, '1', '0', False, False, False, True]	identify barriers

G

Policy when ignoring the person's situation

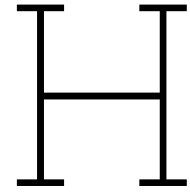
Table G.1: Optimal policy for the model using the first reward function and without considering the person's situation as part of the state. States are of the form [number of changes made to the plan, explaining why planning is useful was done, identifying barriers was done, dealing with barriers was done, showing testimonials from other people was done].

State	Action
[0, True, False, False, False]	changes to plan
[0, False, True, False, False]	changes to plan
[0, False, False, False, True]	changes to plan
[0, False, False, False, False]	changes to plan
[1, True, True, True, True]	changes to plan
[1, True, True, True, False]	changes to plan
[1, True, True, False, True]	changes to plan
[1, True, False, False, True]	changes to plan
[1, True, False, False, False]	show testimonials
[1, False, True, True, True]	changes to plan
[1, False, True, True, False]	changes to plan
[1, False, True, False, False]	deal with barriers
[1, False, False, False, True]	explain planning
[1, False, False, False, False]	identify barriers
[2, True, True, True, False]	show testimonials
[2, True, True, False, True]	deal with barriers
[2, True, False, False, True]	identify barriers
[2, False, True, True, True]	explain planning
[2, False, True, True, False]	explain planning

Since the states only contain the actions already done, this is the only aspect that changes when a persuasive strategy is used. For example, the initial state is the one where no actions have been done ([0, False, False, False, False]), and the policy indicates that making changes to the plan is the optimal action. Then, the state becomes the one where making changes to the plan was done once ([1, False, False, False, False]) and its optimal action is identifying barriers. This forms a chain of predetermined states, actions, and next states, meaning a fixed order of actions. The order defined by the optimal policy is:

1. changes to plan
2. identify barriers
3. deal with barriers

4. changes to plan
5. explain planning
6. show testimonials



Detailed description of the investigation into the effect of the persuasive strategies

Paired Bayesian t -tests were performed on the original data (confidence with values from 0 to 10 and perceived usefulness with values from -10 to 10). For each persuasive strategy, the values of confidence and perceived usefulness before using the persuasive strategy were compared to the values after using the persuasive strategy.

Table H.1 shows the effects of each persuasive strategy on confidence, while table H.2 shows the effects of each persuasive strategy on perceived usefulness. The interpretation of the posterior probabilities is the one from Chechile [26], extended to probabilities below 0.5 based on Andraszewicz et al. [6], with the mention that Andraszewicz et al. use Bayes factors instead of posterior probabilities. The interpretation of Cohen's d is the one from Cohen [27]. In both tables, the mean differences are close to zero, with the 95% highest density intervals of the confidence intervals all containing zero, and with Cohen's d being small in most cases, with a few situations where it is moderate. Looking at the posterior probabilities, all of them are close to 0.5, which means it is not worth betting on there being an increase (or decrease for probabilities smaller than 0.5) in the value of confidence or perceived usefulness for any of the persuasive strategies. Thus, we can conclude that the effect of each persuasive strategy on the person's state is, in general, small.

Since the simulation in Q5 indicated that, over time, people's states do change enough to get to the state with high confidence and perceived usefulness, further exploratory analysis was done, to check the difference in the state at the start and end of the dialogue. Table H.3 shows the results of this analysis. Indeed, the mean of the differences are no longer close to zero, and zero is not included in the 95% HDI CIs. Cohen's d indicates a moderate effect on both confidence and perceived usefulness, and the posterior probabilities suggest a strong difference between the initial and final states. As the posterior probabilities are above 0.999 in both cases, we can make a very strong bet on there being an increase after using all of the persuasive strategies.

Table H.1: Results of paired Bayesian *t*-tests on confidence before and after using a persuasive strategy.

Persuasive strategy	Mean, standard deviation, and 95% HDI CI of the difference	Probability of mean being larger than zero and smaller than zero	Cohen's <i>d</i>
Changes to plan first time	mean -0.008 sd 0.056 95% HDI CI [-0.001, 0.001]	larger 0.487 smaller 0.513	0.32 (moderate)
Changes to plan second time	mean <0.001 sd <0.001 95% HDI CI [-0.00014, 0.00014]	larger 0.498 smaller 0.502	0.11 (small)
Explain planning	mean <0.001 sd <0.001 95% HDI CI [-0.00016, 0.00017]	larger 0.498 smaller 0.502	0.15 (small)
Identify barriers	mean <0.001 sd <0.001 95% HDI CI [-0.00019, 0.00018]	larger 0.501 smaller 0.499	-0.07 (small)
Deal with barriers	mean <0.001 sd <0.001 95% HDI CI [-0.00025, 0.00024]	larger 0.5 smaller 0.5	0.20 (moderate)
Show testimonials	mean <0.001 sd <0.001 95% HDI CI [-0.00027, 0.00025]	larger 0.487 smaller 0.513	0.21 (moderate)

Table H.2: Results of paired Bayesian *t*-tests on perceived usefulness before and after using a persuasive strategy.

Persuasive strategy	Mean, standard deviation, and 95% HDI CI of the difference	Probability of mean being larger than zero and smaller than zero	Cohen's <i>d</i>
Changes to plan first time	mean <0.001 sd <0.001 95% HDI CI [-0.00043, 0.00042]	larger 0.493 smaller 0.507	0.20 (small)
Changes to plan second time	mean <0.001 sd <0.001 95% HDI CI [-0.00024, 0.00022]	larger 0.494 smaller 0.506	0.07 (small)
Explain planning	mean <0.001 sd <0.001 95% HDI CI [-0.00048, 0.00048]	larger 0.499 smaller 0.501	0.15 (small)
Identify barriers	mean <0.001 sd <0.001 95% HDI CI [-0.00024, 0.00025]	larger 0.503 smaller 0.497	-0.003 (small)
Deal with barriers	mean <0.001 sd <0.001 95% HDI CI [-0.00027, 0.00027]	larger 0.5 smaller 0.5	0.11 (small)
Show testimonials	mean <0.001 sd <0.001 95% HDI CI [-0.00032, 0.00030]	larger 0.501 smaller 0.499	0.01 (small)

Table H.3: Results of paired Bayesian *t*-tests at the start and end of the conversation.

Feature	Mean, standard deviation, and 95% HDI CI of the difference	Probability of mean being larger than zero and smaller than zero	Cohen's <i>d</i>
Confidence	mean 1.6 sd 1.9 95% HDI CI [1.2, 2.0]	larger > 0.999 smaller < 0.001	0.75 (moderate)
Perceived usefulness	mean 0.68 sd 1.3 95% HDI CI [0.30, 1.1]	larger > 0.999 smaller < 0.001	0.43 (moderate)



Detailed description of the thematic analysis

Table I.1: Codes for the thematic analysis

Motivating aspects		
Number	Name	Description
1	the plan	the plan itself was motivating
2	tailoring the plan	being able to modify the plan was motivating
3	coping planning	discussing or dealing with barriers was motivating
4	planning on time slots	considering available time was motivating
5	achievable plan	people felt like the plan was achievable and this motivated them
6	considering energy levels	considering how much energy people have was motivating
7	nothing	nothing about the conversation was motivating
8	testimonials	people were motivated by seeing other people plan and achieve goals
9	encouraging to do/be better	people felt like the agent was encouraging them to reach their goals
10	meaningful information	learning about physical activity, goals, and habits motivated people

Demotivating aspects		
Number	Name	Description
1	repeating	people disliked when the virtual coach repeated things it had said previously and this was demotivating for them
2	lack of tailored answers	people felt like the virtual coach was not replying based on what people typed and this was demotivating
3	testimonials	seeing other people succeed was not motivating
4	nothing	nothing about the chat was demotivating
5	lack of deeper understanding	the virtual coach did not properly understand more complex issues and this demotivated people
6	robotic	the virtual coach seemed robotic and this was demotivating
7	difficulty of plan	the plan seemed difficult and this demotivated people
8	long chat	the length of the chat was demotivating
9	not enough plan customisability or detail	not seeing how to do the walks after week 2, or not being able to customise them was demotivating

After reading the motivating and demotivating aspects indicated by people in the responses, a set of initial codes was generated and is presented in Table I.1. Every response was then coded, with the possibility of assigning multiple codes to the same response. A second coder with a background in computer science was then trained on 11 examples of motivating aspects and 11 examples of demotivating aspects and tested on 12 examples of each. Disagreements only occurred for two of the 12 responses to the motivating aspects used for testing and were settled by discussion. Afterwards, the second coder independently coded the rest of the responses, while the first coder double checked that the initially assigned codes conform to the results of the discussion about the disagreements.

To determine which codes were to be kept in the final analysis, we computed both Cohen's Kappa [28] and Brennan-Prediger's Kappa [16]. The results for the motivating factors are shown in Table I.2 and the results for the demotivating factors are shown in Table I.3. According to Cohen's Kappa (interpreted as indicated by Landis and Koch [56]), there is almost perfect agreement for both the motivating and the demotivating factors. For all codes, Cohen's Kappa was above 0.4 (so they had at least moderate agreement) for both motivating and demotivating aspects, so all codes were kept as part of the analysis.

Table I.2: Cohen's Kappa and percent agreement for the entire data set and per code for the motivating factors.

Input	Cohen's Kappa	Brennan-Prediger's Kappa
Entire data set	0.87	0.97
Code 1	0.70	0.87
Code 2	0.54	0.92
Code 3	0.91	0.97
Code 4	1.00	1.00
Code 5	1.00	1.00
Code 6	0.90	0.99
Code 7	1.00	1.00
Code 8	0.97	0.99
Code 9	0.85	0.99
Code 10	0.79	0.97

Table I.3: Cohen's Kappa and percent agreement for the entire data set and per code for the demotivating factors.

Input	Cohen's Kappa	Brennan-Prediger's Kappa
Entire data set	0.95	0.99
Code 1	1.00	1.00
Code 2	0.79	0.99
Code 3	1.00	1.00
Code 4	0.97	0.99
Code 5	0.79	0.97
Code 6	1.00	1.00
Code 7	0.64	0.95
Code 8	1.00	1.00
Code 9	1.00	1.00

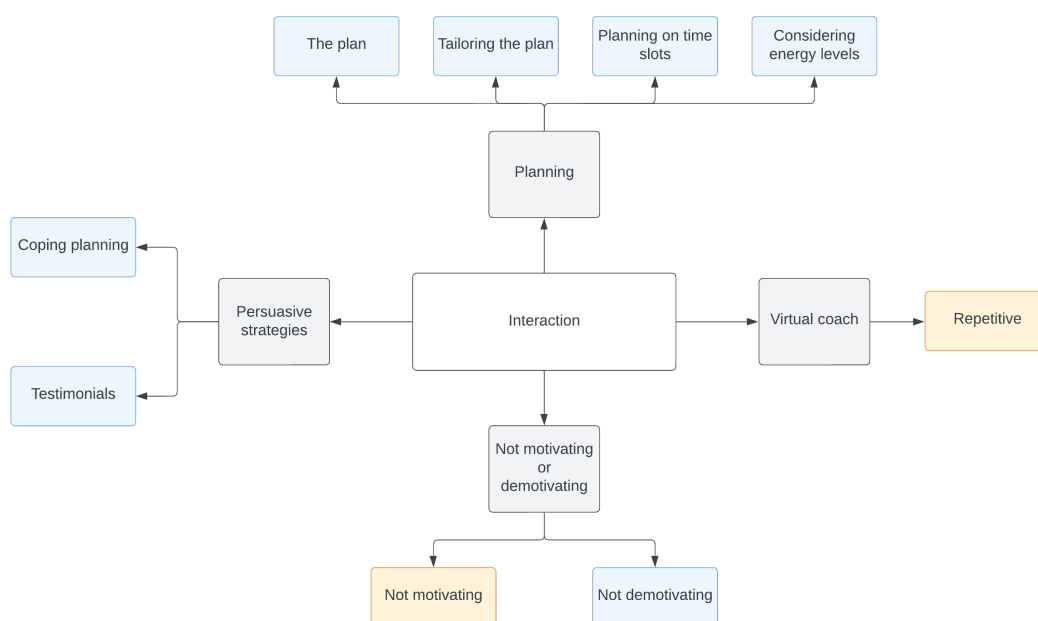


Figure I.1: Themes identified for the motivating (blue) and demotivating (orange) aspects.

The codes were then organised into four themes as shown in Figure I.1. Only the codes which appeared in 9 or more responses are included in the figure and shown through examples in the explanations below.

Persuasive strategies

Coping planning was appreciated, with people indicating that thinking about barriers and of ways to overcome them was something they had not considered previously and was seen as helpful, e.g. participant 21 who said "When Jamie asked me to come with solutions to counteract my obstacles to moving forward with my walking plan."

The testimonials from other people who created plans to reach their goals were regarded positively (e.g. participant 89 "When Jamie brought up instances of people who had followed through with their fitness plan and how they achieved it"), but they were also criticised when they described by some people who felt like the people who gave the testimonials were very different from themselves.

Planning

Considering the time when they were free and the energy levels (e.g. participant 53 "I like that they took interest in what's better for me in relation to my schedules and energy levels during the day") were appreciated by people.

Additionally, having the plan (e.g. participant 113 "When Jamie gave me the full plan") and being able to customise it (e.g. participant 96 "the part where you could easily change your plan if you were not satisfied with what he made") were aspects seen as motivating. Some people also thought that the plan was achievable and that motivated them to follow it.

On the other hand, only being able to see the first two weeks and customise them was seen as an unnecessary restriction, and the fact that the plan stretched quite far into the future made it seem difficult to complete.

Virtual coach

The way in which the virtual coach speaks was appreciated, with some people indicating that the messages provided useful information about planning and encouraged them to stick to their plans.

The main complaint received was that Jamie repeated the three questions for confidence, perceived usefulness and attitude too much (e.g. participant 95 "The repeated questions (3 questions with a rating scale) - they made me feel as I'm not sure of my ratings"), which lead to long dialogues (e.g. participant 46 "The chat was too long, took much time, it was boring").

Furthermore, people indicated that Jamie seems robotic, and talks too much, making the dialogue too long. Another aspect which disappointed people was the lack of tailored answers that Jamie gave, likely due to the lack of a deeper understanding of what the person was saying, as some people pointed out.

Not motivating or demotivating

There were people who indicated that nothing was motivating, e.g. participant 70 "I did not find Jamie to be very motivating," as well as people who indicating that nothing was demotivating, e.g. participant 36 "I didn't find any part of the conversation demotivating, on the contrary".