# Collaboratively Setting Daily Step Goals with a Virtual Coach: Using Reinforcement Learning to Personalize Initial Proposals

Martin Dierikx[1], Nele Albers[1(✉)], Bouke L. Scheltinga[2], and Willem-Paul Brinkman[1]

[1] Intelligent Systems, Delft University of Technology, Delft, Netherlands
{n.albers,w.p.brinkman}@tudelft.nl
[2] Biomedical Signals and Systems, University of Twente, Enschede, Netherlands

**Abstract.** Goal-setting is commonly used in behavior change applications for physical activity. However, for goals to be effective, they need to be tailored to a user's situation (e.g., motivation, progress). One way to obtain such goals is a collaborative process in which a healthcare professional and client set a goal together, thus making use of the professional's expertise and the client's knowledge about their own situation. As healthcare professionals are not always available, we created a dialog with the virtual coach Steph to collaboratively set daily step goals. Since judgments in human decision-making processes are adjusted based on the starting point or anchor, the first step goal proposal Steph makes is likely to influence the user's final goal and self-efficacy. Situational factors impacting physical activity (e.g., motivation, self-efficacy, available time) or how users process information (e.g., mood) may determine which initial proposals are most effective in getting users to reach their underlying previous activity-based recommended step goals. Using data from 117 people interacting with Steph for up to five days, we designed a reinforcement learning algorithm that considers users' current and future situations when choosing an initial step goal proposal. Our simulations show that initial step goal proposals matter: choosing optimal ones based on this algorithm could make it more likely that people move to a situation with high motivation, high self-efficacy, and a favorable daily context. Then, they are more likely to achieve, but also to overachieve, their underlying recommended step goals. Our dataset is publicly available.

**Keywords:** Physical activity · Behavior change · Reinforcement learning · Conversational agent · Goal-setting

## 1 Introduction

Goal-setting is commonly used in behavior change applications for physical activity (e.g., [3,6,39,44,50]). It helps to stay focused on a desired outcome, spend effort toward that outcome, and find effective strategies [37]. However, for goals


N. Baghaei et al. (Eds.): PERSUASIVE 2024, LNCS 14636, pp. 100–115, 2024.
https://doi.org/10.1007/978-3-031-58226-4_9

to be effective, it is recommended that they satisfy criteria such as being specific, measurable, achievable, relevant, and time-bound (SMART) [21] as well as being (re-)evaluated [5,37]. Current physical activity applications commonly do not satisfy these criteria, especially when it comes to tailoring the goal difficulty to a user's ability and re-evaluating goals based on the user's progress [5].

In a traditional offline setting, one way to set goals satisfying these criteria is for a client and healthcare professional to agree on a goal in a collaborative process [10]. Involving clients in the goal-setting process can not only increase their self-efficacy [37], but it also offers the opportunity to combine the expertise of the healthcare professional and the client's knowledge about their own situation. Such collaborative goal-setting has, for example, been recommended for people with diabetes [14] and been applied in the context of asthma management [47]. Since virtual coaches can take the role of such a healthcare professional in eHealth applications for behavior change, we thus wanted to design a collaborative goal-setting dialog with a virtual coach. Besides providing guidance where traditionally healthcare professionals would have, virtual coaches also have the potential to combat the low adherence common to eHealth applications for behavior change [7,26] by fostering engagement, discussing relevant and timely information, showing understanding, and connecting with people [2,27,38].

When designing such a collaborative goal-setting dialog, attention needs to be paid to the starting point. This is because, in a human decision-making process, judgments are commonly made based on the starting point of the process (i.e., an anchor) [52]. All subsequent judgments are then made by adjusting away from that anchor. For example, the first offer in a negotiation has been shown to be a strong predictor of the settling price for purchasing a pharmaceutical plant or the assigning bonus for a new employee [25]. And anchoring values have also been shown to affect self-efficacy in a problem-solving task [12]. This means that the first goal option that is discussed in the goal-setting dialog is likely to influence both the final goal the virtual coach and the user set and the user's self-efficacy regarding achieving that goal.

Since a suitable physical activity goal for the user depends on their current situation (e.g., previous physical activity, depressive symptoms, self-efficacy [42]), the starting point of the goal-setting process should hence also be adapted to a user's current situation. Previous work has, for example, adapted physical activity goals based on a user's routine [11], previous performance [33], or location, step count variation, number of app screens yesterday, and past push notifications [36]. Moreover, since we want to adapt initial goal proposals rather than fixed goals, factors that influence how people process information can also play a role. For instance, a user's mood may influence the degree of message elaboration [8]. When setting multiple short-term (e.g., daily) goals over an extended period of time, however, it is not only the *current* situation of the user that matters but also the *future* one. For example, while setting higher physical activity goals may result in higher physical activity levels, it may also lead to lower goal achievement [13] and thus potentially lower engagement in the future [49]. Setting initially lower goals, on the other hand, may allow people to make small wins and thus increase their motivation [4].

One framework that allows us to consider both current and future user situations (i.e., states) is Reinforcement Learning (RL) [48]. RL, with a consideration of current and future states, has previously been applied to adapt weekly step goals to people's previous activity and self-efficacy [59] or determine when to send physical activity notifications [53]. Here, we investigate whether RL is also useful when choosing initial goal proposals in a collaborative dialog for physical activity. To this end, we conducted a study in which 117 people interacted with the text-based virtual coach Steph for up to five days. Each day, participants and Steph collaboratively set a daily goal for walking, which is easily accessible to most people [34], has documented health benefits [34], and is one of the easiest and most acceptable forms of physical activity since it can be integrated into everyday life [15]. In this collaboration, Steph first determined people's current situation with regard to mood, sleep quality, available time, motivation, and self-efficacy before giving a step goal proposal that could be iteratively refined afterward. The proposal could thereby take five different forms, each based on adjusting an underlying previous activity-based recommended goal in a different way (i.e., ± 0, 200, or 400 steps). On the next day, Steph asked the user about the number of steps they took the previous day before initiating the setting of a new goal. Based on this study's data, this paper's contribution is threefold. First, we provide insights into the effects of initial goal proposals in a collaborative goal-setting dialog. Second, we contribute an RL model that optimizes the choice of initial proposals based on people's current and future states. And third, we publish our dataset to aid future work on adaptive collaborative goal-setting.

## 2 Materials and Methods

### 2.1 Virtual Coach

We developed the text-based virtual coach Steph [18] in Rasa [9]. Steph introduced itself as helping people set daily step goals toward the ultimate goal of taking 10,000 steps every day [55]. In each session, Steph asked about people's current state based on their mood, sleep quality, available time, motivation, and self-efficacy. Afterward, Steph computed a recommended daily step goal based on the user's previous walking behavior using the percentile algorithm by Adams et al. [1][1]. Based on this recommended step goal, Steph gave the user three goal options, each 100 steps apart, as well as the possibility to indicate that they wanted a different goal. This way, users were given a say in determining their goal and nudged toward picking one of the three options. If users indicated wanting a higher goal, Steph congratulated them for wanting to challenge themselves but also warned them that taking too many steps might lead to injuries; if users

[1] We rounded the resulting number to the nearest 100 since people tend to need more time to process non-rounded numbers [31] and to put extra effort into completing a rounded goal if they are close to it [43].

indicated wanting a lower goal, Steph expressed understanding but said that taking at least some steps is good for them. Users could indicate wanting a lower or higher goal up to five times with a minimum goal of 2,000 and a maximum goal of 10,000 steps[2]. Once the user had decided on a goal, Steph congratulated them on their choice, gave a few examples of how to easily take steps during the day, and sent a reminder message with the goal on Prolific. The next session started by asking users about the number of steps they took on the previous day. In its dialog style, Steph followed principles from motivational interviewing such as expressing empathy and acknowledging answers [51]. As part of a social communication style, Steph further used informal language (e.g., "Aww, that's annoying") and emojis, made use of positively valenced words (e.g. "great", "cool"), and reacted enthusiastically to users' inputs. Such a social communication style has, for example, been shown to increase customer satisfaction with chatbots [58]. A demo video of the dialog can be found online [16].

### 2.2 Personalizing Initial Step Goal Options

We can define our approach as a Markov Decision Process (MDP) $\langle S, A, R, T, \gamma \rangle$. The action space $A$ consisted of five ways of personalizing the step goal proposals, the reward function $R : S \times A \rightarrow \mathbb{R}$ was determined by the difference between the recommended step goal and the number of steps a person took, $T : S \times A \times S \rightarrow [0, 1]$ was the transition function, and the discount factor $\gamma$ was set to 0.85 to favor rewards obtained earlier over rewards obtained later. The finite state space $S$ described the state a user was in and was captured by their mood, sleep quality, available time, motivation, and self-efficacy. The goal of an agent in an MDP is to learn an optimal policy $\pi^* : S \rightarrow \Pi(A)$ that maximizes the expected cumulative discounted reward $\mathbb{E}\big[\sum_t^\infty \gamma^t r_t\big]$ for acting in the environment. The optimal Q-value function $Q^* : S \times A \rightarrow \mathbb{R}$ describes the expected cumulative discounted reward for executing $a$ in state $s$ and $\pi^*$ in all subsequent states. Figure 1 shows how our RL approach is embedded in the goal-setting dialog.

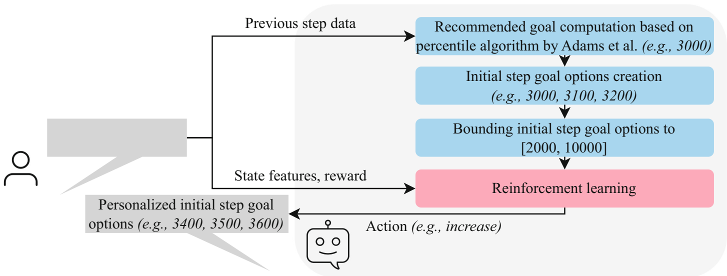



**Fig. 1.** Pipeline for arriving at personalized initial step goal options.

[2] Few people take less than 2,000 daily steps [22,28,56], and taking more than 10,000 steps does not offer much further benefit while making injuries more likely [35,41].

*State Space.* In each session, users provided answers to questions about their mood valence based on the adjectives by Russell [46] mapped to a valence score using the emotion wheel by Kollias et al. [32], sleep quality from the previous night based on Dzierzewski et al. [23], available time and motivation based on the physical activity barrier descriptions by Robbins et al. [45], and self-efficacy based on the definition by Park and Kim [42]. The latter four variables were all measured on 11-point scales from 0 to 10. Mood valence, sleep quality, and available time were subsequently added to create a "daily context"-state that described factors that could influence the effectiveness of the actions but were unlikely to also be affected by the actions. This "daily context"-state could take values between 0 and 30. The intuition is that the higher the value, the more favorable is the daily context for physical activity.

*Actions.* Our action space consisted of five ways of personalizing the initial step goal options: *decrease*, *slightly decrease*, *do not change*, *slightly increase*, and *increase*. *Slightly decrease* and *slightly increase* decreased or increased the three initial goal options by 200 steps, *decrease* and *increase* changed the initial options by 400 steps, and *do not change* kept the original options.

*Reward.* The reward $r$ was based on the absolute difference $\Delta$ between the recommended goal $G$ and the actual number of steps a user took (*Steps*). Taking more steps than recommended is penalized half as much as taking fewer steps:

$$r = \begin{cases} 1 - \frac{\Delta}{G} & if\ Steps \leq G \\ 1 - \frac{\Delta}{2G} & if\ Steps > G. \end{cases}$$

The intuition behind this reward signal is that we primarily want users to reach the recommended number of steps, but also penalize taking too many steps to some degree since this can lead to injuries.

### 2.3 Data Collection and Model Training

*Study.* To collect training data for our algorithm, we conducted an observational study on the crowdsourcing platform Prolific in June and July 2023. We preregistered the study in the Open Science Framework (OSF) [19]. Since a pilot study with 34 people did not result in major changes other than the addition of one question to the post-questionnaire, the data from the pilot study was used as well. The Human Research Ethics Committee of Delft University of Technology approved our study (Letter of Approval number: 3016). Eligible were people who indicated being fluent in English, being between 18 and 65 years old, engaging in physical exercise for at most 150 min per week, having taken no more than 9,000 steps on average per day in the last week, not participating in a physical activity program, having a low risk of getting injured because of walking according to the Physical Activity Readiness Questionnaire 2023 [54], being contemplating or preparing to become more physically active, and having a way to track their steps. Moreover, we used the quality measures on Prolific to choose people who had completed at least one previous study and had an approval rate of at least 90%. Participants further had to live in one of five time zones (GMT, GMT+1, GMT+2, GMT+3, or GMT+4) to ensure that the daily step goals were set in the morning. The way the step goal options were personalized was chosen randomly in all five sessions. 235 people were invited to the first session and 77 people successfully completed all five sessions. Of the 117 participants with at least one transition sample, 60 (51%) were female, 55 (47%) were male, and 2 (2%) indicated another gender. The age ranged from 18 to 56 ($M$ = 28, $SD$ = 8) years. Participants most commonly reported using an iPhone health app (32%), the Samsung Health app (31%), or a smartwatch (24%) to track their steps. The average number of steps per day before the study ranged from 30 to 9,000 ($M$ = 4,402, $SD$ = 2,383). Participants who successfully completed a study part were paid based on the minimum payment rules on Prolific (i.e., six GBP per hour). Participants were informed that their payment was independent of their achieving their step goals.

*Collected Data.* We collected 381 $\langle s, a, r, s' \rangle$-samples from 117 people, where $s$ is the state, $a$ the action, $r$ the reward, and $s'$ the next state. Across all 381 samples, the initially proposed step goals were rejected a total of 100 times ($M$ = 0.26, $SD$ = 0.67). People reached their goals 66% of the time and found their goals relatively easy to reach. Specifically, the mean goal difficulty rating provided after the five sessions was 2.04 ($SD$ = 2.06), rated on a scale from −5 ("It was very difficult to reach the daily goals") to 5 ("It was very easy to reach the daily goals"). The average number of steps taken per day before the study was 4,549 ($SD$ = 4,350) for the 75 people who completed the post-questionnaire. On the fifth and final day of the study, these people took an average of 5,367 steps ($SD$ = 3,353). Based on a paired Bayesian $t$-test, this corresponds to a mean increase of 1,087 steps ($SD$ = 1,298, 95%-HDI = [710, 1,478]). Furthermore, we collected data on people's experiences of their interaction with Steph after the five sessions using the short form of the Artificial Social Agent (ASA) Questionnaire [24]. We obtained a mean score of 19.32, which is higher than the scores of 9 of the 14 agents tested by Fitrianie et al. [24].

*State Space Reduction.* To reduce the size of the state space and thus the amount of required data, we transformed the three state features (motivation, self-efficacy, and the "daily context"-state) into binary features based on whether a value was greater than or equal to the median (1) or less than the median (0). The final state space thus had size $|S| = 2^3 = 8$. We refer to states with binary strings such as 001 (here motivation and self-efficacy are 0 and the "daily context"-state is 1).

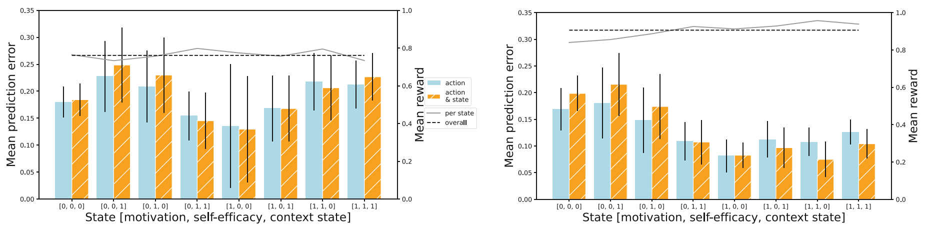
*Model Training.* Using the reward and transition functions estimated from the data, we computed $Q^*$ using value iteration. Since some states were much more common than others, we had fewer than ten samples for some state-action combinations. To reduce overfitting, we added samples with the overall mean reward for the reward prediction for those state-action combinations with few samples. Similarly, we balanced with an equal probability of all next states when estimating the transition function. Overall, we imputed 125 samples.

## 3 Results

We now investigate each of our analysis questions. For each of them, we first describe our setup, followed by our findings and the resulting answer to the question. Our data and analysis code are available online [20].

**AQ1: How well do states predict behavior after proposing personalized step goals?**

*Setup.* Knowing the state a user is in may help to predict their behavior after using different ways of personalizing their initial step goals (i.e., actions). The behavior in our case is how close people get to their underlying recommended step goal, which is captured by the reward function. We compared two approaches for predicting the reward: 1) the mean reward per action, and 2) the mean reward per action and state. We used leave-one-out cross-validation for the 117 participants with at least one transition sample to compare the two approaches based on the mean $L_1$-error and its Bayesian 95% credible interval (CI) [40] per state. In contrast to the often used frequentist confidence intervals, Bayesian CIs provide information on the most likely values (i.e., a likely range) [29]. If the mean of one of the approaches is outside of the credible interval of the other approach, we regard this as a credible indication that values are different.




(a) Original reward. (b) Without overachieving penalty.

**Fig. 2.** Left axis: Mean $L_1$-error with 95%-CIs for predicting rewards based on 1) the mean reward per action and 2) the mean reward per action and state. Right axis: Mean reward overall and per state.

*Results.* Figure 2a shows that none of the two approaches for predicting the reward clearly performs better for any of the eight states. The mean rewards per state also are very similar, ranging from 0.73 for state 001 to 0.80 for state 011. However, if, for exploratory purposes, we modify the original reward by removing the penalty for overachieving (i.e., taking more steps than the recommended goal), considering the state a user is in does improve the reward prediction in some states (Fig. 2b). The mean modified reward also differs more clearly between states, with the mean modified reward being generally higher when more state features are high. This can be explained by the observation that even though the mean rewards are similar in the eight states (Fig. 2a), the underlying behaviors differ. Specifically, many people *under*achieve their recommended goals in states with low state features, whereas many people *over*achieve their goals in states with high state features (Fig. 3). Thus, more people reach their recommended goals in states with high state features.

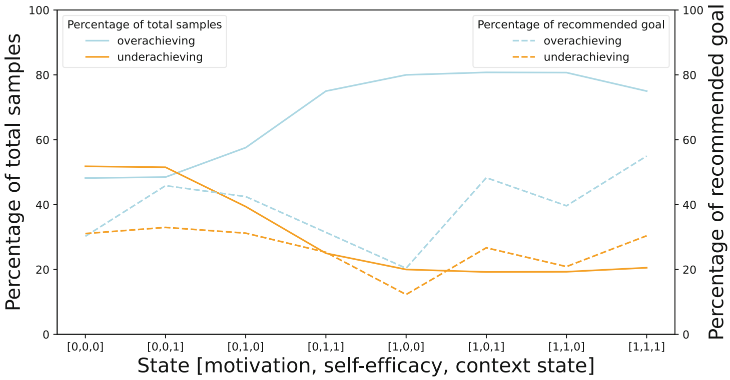



**Fig. 3.** Left axis (*continuous lines*): percentage of samples where more or fewer steps were taken than the recommended goal per state. Right axis (*dashed lines*): percentage of the recommended goal that was over- or underachieved.

*Answer to AQ1.* Knowing a user's state does not provide a clear benefit for predicting their behavior after proposing personalized step goals. If, however, we remove the penalty for taking more steps than one's recommended goal, knowing a user's state does offer a benefit for some states. So states do matter when it comes to predicting whether people reach their recommended goal.

**AQ2: How well do states predict next states after proposing personalized step goals?**

*Setup.* By making a personalized initial proposal, we would ideally want people to move to a next state in which they are very likely to reach their recommended step goal and thus make more progress toward the long-term goal of 10,000 steps per day. Thus, we need to be able to predict the state after proposing personalized step goals. We again used leave-one-out cross-validation to compare three ways of predicting the next states for the samples from the left-out person: 1) assigning an equal probability to all states, 2) predicting that people stay in their current state, and 3) using the transition function estimated from the training data. We compared the three approaches based on the mean likelihood of the next state and its 95%-CI per state. A higher likelihood suggests that next states can be predicted better. Again, if the mean of one of the approaches lies outside of the credible interval of another approach, we regard this as a credible indication that values are different.

*Results.* Figure 4 shows that considering the current state, either by predicting that people stay in their current state or by using the estimated transition function to predict next states, generally helps to predict the next state. This suggests that state transitions do not occur uniformly at random. In two states, namely, 000 and 111, predicting that people stay in their current state leads to the highest mean likelihood of next states. In both these states, the corresponding means are clearly outside the 95%-CIs of the other two approaches, suggesting that the values are different. This shows the high probability of staying in these two states, which are states in which people either commonly underachieve their recommended goals or overachieve them by a lot (Fig. 3).

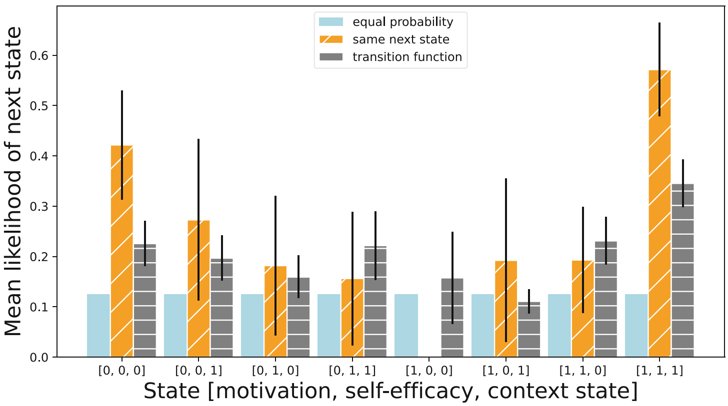


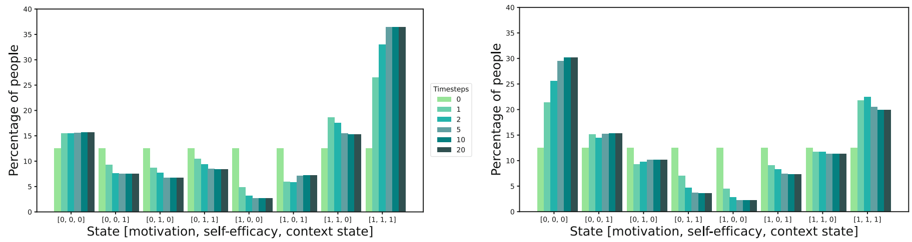

**Fig. 4.** Comparison of three approaches to predicting next states with regard to the mean likelihood of next states with 95%-CIs for each state.

*Answer to AQ2.* Our results show that considering the current state a user is in helps to predict their next state after proposing personalized step goals. Furthermore, once people are in states with low or high values for all three state features, they tend to stay there. This suggests that it is difficult to move people out of the state where more than half of them do not reach their recommended goal (i.e., state 000). However, once people are in a state where almost 80% reach or overachieve their goal (i.e., state 111), they tend to stay there.

**AQ3: What is the effect of (multiple) optimal step goal proposals on users' states?**

*Setup.* We would like users to ultimately move to states in which they are most likely to reach their recommended goals. Starting from an equal distribution of 8,000 simulated people across the states, we calculated the percentage of people in each state after following the optimal policy $\pi^*$ for a certain number of time steps.

*Results.* Figure 5a shows that after 20 time steps, the largest percentage of people (36.5%) is in state 111, which is a state in which most people reach or overachieve their goal. However, the number of people in state 000, which is the state where most people do not reach their goal, also slightly increases to 15.7%.




(a) (b)

**Fig. 5.** Percentage of people in each state after following (a) the optimal policy $\pi^*$ and (b) the worst policy $\pi^-$ for varying numbers of time steps.

*Answer to AQ3.* While choosing the optimal way of personalizing step goals multiple times allows most people to move to and stay in states where they are most likely to reach or overachieve their recommended goal, a considerable number of people also remain in the state in which they are least likely to reach their recommended goal.

**AQ4: How do optimal and sub-optimal step goal proposals compare in their effects on behavior?**

*Setup.* So far, we are, to some degree, able to predict next states and choose sequences of personalized step goal proposals that move people to better states. However, how much does the choice of step goal proposal personalization matter? To examine this, we calculated the mean reward per transition over time when following 1) the optimal policy $\pi^*$, 2) the worst policy $\pi^-$, and 3) the average policy $\pi^\sim$. $\pi^\sim$ is a theoretical policy for comparison purposes in which each action is taken $\frac{1}{|A|}$ times for each person at each time step, where $|A|$ is the number of actions. We simulated 117 people, initially distributed across the states as in the first session of our study.

*Results.* The mean reward per transition for $\pi^*$ is 0.11 (15.8%) higher than for $\pi^-$ and 0.06 (8.0%) higher than for $\pi^\sim$ after 100 time steps (Fig. 6). This suggests that when it comes to obtaining a higher reward, the choice of goal proposal personalization is not so important. If, however, we again remove the penalty for overachieving (i.e., only consider whether people reach their recommended goal), the choice of goal proposal personalization does matter. This can be seen by comparing Fig. 5b to Fig. 5a. Specifically, following the worst policy leads to more people in state 000 and fewer people in state 111 after 20 simulated time steps, which means that fewer people reach their recommended goals.

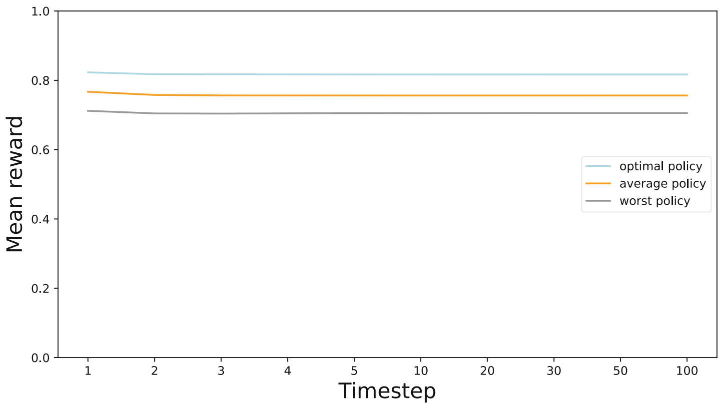



**Fig. 6.** Mean reward per transition over time for three policies.

*Answer to AQ4.* If we want people to achieve their goal and penalize both under- and overachieving, personalizing step goal proposals optimally poses an advantage of between 8.0% and 15.8% compared to doing so in an average or the worst possible way. Following the worst policy does, however, cause people to more frequently move to states in which they are more likely to underachieve. Thus, if we do not mind people overachieving their goals and just want people to reach their goals, optimally personalizing step goal proposals does matter.

## 4 Discussion and Conclusion

This work presented a dialog to collaboratively set daily step goals with a virtual coach and specifically examined the effect of initial step goal proposals that are created by adjusting an underlying previous activity-based recommended goal in one of five possible ways (i.e., ± 0, 200, or 400). We find that user states based on their mood, sleep quality, available time, motivation, and self-efficacy help to predict whether people reach their recommended goal after these proposals are made (*AQ1*). Moreover, these current user states are also predictive of the user states on the next day and thus of whether people reach their recommended goals in the future (*AQ2*). Regarding long-term effects, our simulations show that choosing optimal step goal proposals based on an RL algorithm that considers people's current and future states allows most people to move to and stay in states where they are very likely to reach their recommended goal (*AQ3*). However, some people always remain in the state in which they are least likely to reach their recommended goal. Our simulations further show that it matters which initial step goal proposal the virtual coach makes (*AQ4*). Specifically, more people move to good states (i.e., states where more people reach their recommended goal) if an initial proposal that is optimal based on the RL algorithm is made than in the case of a sub-optimal proposal.

While the likelihood of users reaching their recommended goal differs between user states, this is not the case for our originally devised reward signal that also penalizes taking more steps than one's recommended goal. This is because people in good states often not only reach their recommended goal but substantially exceed it, which also means that our RL algorithm over time tends to move people to states where they overachieve the number of steps they are recommended to take. We had originally decided to penalize overachieving because of the risk of injuries that exists especially for physically inactive people like our participants due to low strength and cardiovascular fitness levels [30]. However, it is not yet sufficiently clear at which point the injury risk outweighs the benefits of more physical activity, particularly because this also depends on the walking speed [30]. Instead of penalizing overachieving in the reward signal, it may be better to more thoroughly educate users on the potential risks of exceeding recommended step goals (e.g., by reminding overachieving users of injury risks and how they can become more physically active in a healthy way).

Besides the handling of overachieving recommended goals, our study has several further limitations. First, due to the high cost of collecting human data like ours, we obtained a relatively limited dataset of 381 samples. We thus turned our state features into binary features, but we still had relatively few samples for some states. It would be interesting to repeat our analysis with more data. To facilitate this, we have made our dataset publicly available [20]. In addition, while we did gather data from human subjects, we did not study the actual long-term effects of making optimal initial step goal proposals based on our RL algorithm. Moreover, even though our participants were informed that their payment for completing the daily goal-setting dialogs was independent of their achieving their step goals, they might have felt at least some obligation to take daily steps. Future work should compare the long-term effects of different ways of choosing initial step goal proposals in the wild. Lastly, there might be other factors that influence the effectiveness of different initial step goal proposals, such as people's social environment [42] and their traits (e.g., personality [57]).

In summary, we have created a virtual coach to collaboratively set daily step goals with users and an RL algorithm that adapts initial step goal proposals based on current and future user states. Simulations show that initial proposals matter: using RL to optimize initial step goal proposals could make it more likely that people move to a state with high motivation, high self-efficacy, and a favorable daily context. In such a state, people are more likely to reach, but also to exceed, their recommended number of steps. Designers of collaborative goal-setting dialogs should thus choose the first proposal carefully and based on users' situations while accounting for the possibility that users exceed their physical activity recommendations.

**Acknowledgments.** This work is based on the publicly available Master's thesis by Dierikx [17]. The Master's thesis was written as 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. The authors acknowledge the help they received from Kristell Penfornis and Anke Versluis in designing the goal-setting dialog.

## References

1. Adams, M.A., Sallis, J.F., Norman, G.J., Hovell, M.F., Hekler, E.B., Perata, E.: An adaptive physical activity intervention for overweight adults: a randomized controlled trial. PLoS ONE **8**(12), e82901 (2013). https://doi.org/10.1371/journal.pone.0082901
2. Aggarwal, A., Tam, C.C., Wu, D., Li, X., Qiao, S.: Artificial intelligence-based chatbots for promoting health behavioral changes: systematic review. J. Med. Internet Res. **25**, e40789 (2023). https://doi.org/10.2196/40789
3. Albers, N., Hizli, B., Scheltinga, B.L., Meijer, E., Brinkman, W.P.: Setting physical activity goals with a virtual coach: vicarious experiences, personalization and acceptance. J. Med. Syst. **47**(1), 15 (2023). https://doi.org/10.1007/s10916-022-01899-9
4. Amabile, T., Kramer, S.: The Progress Principle: Using Small Wins to Ignite Joy, Engagement, and Creativity at Work. Harvard Business Press, Brighton (2011)
5. Baretta, D., Bondaronek, P., Direito, A., Steca, P.: Implementation of the goal-setting components in popular physical activity apps: review and content analysis. Digit. Health **5**, 2055207619862706 (2019). https://doi.org/10.1177/2055207619862706
6. Beinema, T., op den Akker, H., van Velsen, L., Hermens, H.: Tailoring coaching strategies to users' motivation in a multi-agent health coaching application. Comput. Hum. Behav. **121**, 106787 (2021). https://doi.org/10.1016/j.chb.2021.106787
7. Beun, R., et al.: Improving adherence in automated e-coaching - a case from insomnia therapy. In: Meschtscherjakov, A., de Ruyter, B.E.R., Fuchsberger, V., Murer, M., Tscheligi, M. (eds.) PERSUASIVE 2016. LNCS, vol. 9638, pp. 276–287. Springer, Heidelberg (2016). https://doi.org/10.1007/978-3-319-31510-2_24
8. Bless, H., Bohner, G., Schwarz, N., Strack, F.: Mood and persuasion: a cognitive response analysis. Pers. Soc. Psychol. Bull. **16**(2), 331–345 (1990). https://doi.org/10.1177/0146167290162013
9. Bocklisch, T., Faulkner, J., Pawlowski, N., Nichol, A.: Rasa: open source language understanding and dialogue management. arXiv preprint arXiv:1712.05181 (2017)
10. Bodenheimer, T., Handley, M.A.: Goal-setting for behavior change in primary care: an exploration and status report. Patient Educ. Couns. **76**(2), 174–180 (2009). https://doi.org/10.1016/j.pec.2009.06.001
11. Cabrita, M., op den Akker, H., Achterkamp, R., Hermens, H.J., Vollenbroek-Hutten, M.M.R.: Automated personalized goal-setting in an activity coaching application. In: Postolache, O., van Sinderen, M., Ali, F.H., Benavente-Peces, C. (eds.) SENSORNETS 2014 - Proceedings of the 3rd International Conference on Sensor Networks, Lisbon, Portugal, 7–9 January 2014, pp. 389–396 (2014)
12. Cervone, D., Peake, P.K.: Anchoring, efficacy, and action: the influence of judgmental heuristics on self-efficacy judgments and behavior. J. Pers. Soc. Psychol. **50**(3), 492–501 (1986)
13. Chevance, G., et al.: Goal setting and achievement for walking: a series of n-of-1 digital interventions. Health Psychol. **40**(1), 30–39 (2021)

14. Davis, J., et al.: 2022 national standards for diabetes self-management education and support. Sci. Diab. Self-Manag. Care **48**(1), 44–59 (2022). https://doi.org/10.1177/26350106211072203
15. Department of Health and Social Care: UK chief medical officers' physical activity guidelines (2019). https://assets.publishing.service.gov.uk/media/5d839543ed915d52428dc134/uk-chief-medical-officers-physical-activity-guidelines.pdf. Accessed 13 Nov 2023
16. Dierikx, M.: Example dialog with the virtual coach steph for setting daily step goals (2023). https://youtu.be/FSpG-G0zc-o
17. Dierikx, M.: Using reinforcement learning to personalize daily step goals for a collaborative dialogue with a virtual coach. Master's thesis (2023). http://resolver.tudelft.nl/uuid:4f2c12de-9b9f-4e3f-ad3a-902947d693bb
18. Dierikx, M.: Using Reinforcement Learning to Personalize Daily Step Goals for a Collaborative Dialogue with a Virtual Coach: code for the virtual coach (2023). https://doi.org/10.5281/zenodo.8382413
19. Dierikx, M., Albers, N., Brinkman, W.P.: Daily collaborative personalized step goal-setting with a virtual coach (2023). https://doi.org/10.17605/OSF.IO/6JQPK
20. Dierikx, M., Albers, N., Scheltinga, B.L., Brinkman, W.P.: Collaboratively setting daily step goals with a virtual coach: using reinforcement learning to personalize initial proposals - data and analysis code (2023). https://doi.org/10.4121/53f2d238-77fc-4045-89a9-fb7fa2871f1d
21. Doran, G.T.: There's a smart way to write management's goals and objectives. Manage. Rev. **70**(11), 35–36 (1981)
22. Dwyer, T., et al.: The inverse relationship between number of steps per day and obesity in a population-based sample-the ausdiab study. Int. J. Obes. **31**(5), 797–804 (2007)
23. Dzierzewski, J.M., et al.: Exercise and sleep in community-dwelling older adults: evidence for a reciprocal relationship. J. Sleep Res. **23**(1), 61–68 (2014). https://doi.org/10.1111/jsr.12078
24. Fitrianie, S., Bruijnes, M., Li, F., Abdulrahman, A., Brinkman, W.P.: The artificial-social-agent questionnaire: establishing the long and short questionnaire versions. In: Martinho, C., Dias, J., Campos, J., Heylen, D. (eds.) Proceedings of the 22nd ACM International Conference on Intelligent Virtual Agents, IVA 2022. Association for Computing Machinery, New York (2022)
25. Galinsky, A.D., Mussweiler, T.: First offers as anchors: the role of perspective-taking and negotiator focus. J. Pers. Soc. Psychol. **81**(4), 657–669 (2001). https://doi.org/10.1037/0022-3514.81.4.657
26. Greenhalgh, T., et al.: Beyond adoption: a new framework for theorizing and evaluating nonadoption, abandonment, and challenges to the scale-up, spread, and sustainability of health and care technologies. J. Med. Internet Res. **19**(11) (2017). https://doi.org/10.2196/jmir.8775
27. He, L., Balaji, D., Wiers, R.W., Antheunis, M.L., Krahmer, E.: Effectiveness and acceptability of conversational agents for smoking cessation: a systematic review and meta-analysis. Nicotine Tobacco Res. **25**(7), 1241–1250 (2022)
28. Hirvensalo, M., et al.: Daily steps among finnish adults: variation by age, sex, and socioeconomic position. Scand. J. Public Health **39**(7), 669–677 (2011). https://doi.org/10.1177/1403494811420324
29. Hoekstra, R., Morey, R.D., Rouder, J.N., Wagenmakers, E.J.: Robust misinterpretation of confidence intervals. Psychon. Bull. Rev. **21**, 1157–1164 (2014). https://doi.org/10.3758/s13423-013-0572-3

30. Hootman, J.M., Macera, C.A., Ainsworth, B.E., Addy, C.L., Martin, M., Blair, S.N.: Epidemiology of musculoskeletal injuries among sedentary and physically active adults. Med. Sci. Sports Exerc. **34**(5), 838–844 (2002)
31. Jain, G., Gaeth, G.J., Nayakankuppam, D., Levin, I.P.: Revisiting attribute framing: the impact of number roundedness on framing. Organ. Behav. Hum. Decis. Process. **161**, 109–119 (2020)
32. Kollias, D., et al.: Deep affect prediction in-the-wild: aff-wild database and challenge, deep architectures, and beyond. Int. J. Comput. Vision **127**(6), 907–929 (2019). https://doi.org/10.1007/S11263-019-01158-4
33. Konrad, A., et al.: Finding the adaptive sweet spot: balancing compliance and achievement in automated stress reduction. In: Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, CHI 2015, pp. 3829–3838. Association for Computing Machinery, New York (2015)
34. Lee, I.M., Buchner, D.M.: The importance of walking to public health. Med. Sci. Sports Exerc. **40**(7), S512–S518 (2008)
35. Lee, I.M., Shiroma, E.J., Kamada, M., Bassett, D.R., Matthews, C.E., Buring, J.E.: Association of step volume and intensity with all-cause mortality in older women. JAMA Intern. Med. **179**(8), 1105–1112 (2019)
36. Liao, P., Greenewald, K.H., Klasnja, P.V., Murphy, S.A.: Personalized heartsteps: a reinforcement learning algorithm for optimizing physical activity. Proc. ACM Interact. Mob. Wearable Ubiquitous Technol. **4**(1), 18:1–18:22 (2020)
37. Locke, E.A., Latham, G.P.: Building a practically useful theory of goal setting and task motivation: a 35-year odyssey. Am. Psychol. **57**(9), 705–717 (2002)
38. Montenegro, J.L.Z., da Costa, C.A., da Rosa Righi, R.: Survey of conversational agents in health. Expert Syst. Appl. **129**, 56–67 (2019)
39. Muellmann, S., Forberger, S., Möllers, T., Bröring, E., Zeeb, H., Pischke, C.R.: Effectiveness of ehealth interventions for the promotion of physical activity in older adults: a systematic review. Prev. Med. **108**, 93–110 (2018)
40. Oliphant, T.E.: A bayesian perspective on estimating mean, variance, and standard-deviation from data (2006)
41. Paluch, A.E., et al.: Daily steps and all-cause mortality: a meta-analysis of 15 international cohorts. Lancet Public Health **7**(3), e219–e228 (2022). https://doi.org/10.1016/S2468-2667(21)00302-9
42. Park, H., Kim, N.: Predicting factors of physical activity in adolescents: a systematic review. Asian Nurs. Res. **2**(2), 113–128 (2008)
43. Pope, D., Simonsohn, U.: Round numbers as goals: evidence from baseball, sat takers, and the lab. Psychol. Sci. **22**(1), 71–79 (2011)
44. Rhodes, A., Smith, A.D., Chadwick, P., Croker, H., Llewellyn, C.H., et al.: Exclusively digital health interventions targeting diet, physical activity, and weight gain in pregnant women: systematic review and meta-analysis. JMIR Mhealth Uhealth **8**(7), e18255 (2020). https://doi.org/10.2196/18255
45. Robbins, L.B., Pender, N.J., Kazanis, A.S.: Barriers to physical activity perceived by adolescent girls. J. Midwifery Women's Health **48**(3), 206–212 (2003)
46. Russell, J.A.: A circumplex model of affect. J. Pers. Soc. Psychol. **39**(6), 1161–1178 (1980)
47. Smith, L., et al.: The contribution of goal specificity to goal achievement in collaborative goal setting for the management of asthma. Res. Social Adm. Pharm. **9**(6), 918–929 (2013). https://doi.org/10.1016/j.sapharm.2013.02.002
48. Sutton, R.S., Barto, A.G.: Reinforcement learning: An introduction. Bradford Books (2018)

49. Swann, C., Rosenbaum, S., Lawrence, A., Vella, S.A., McEwan, D., Ekkekakis, P.: Updating goal-setting theory in physical activity promotion: a critical conceptual review. Health Psychol. Rev. **15**(1), 34–50 (2021)
50. Tong, H.L., et al.: A personalized mobile app for physical activity: an experimental mixed-methods study. Digital Health **8**, 20552076221115016 (2022)
51. Treasure, J.: Motivational interviewing. Adv. Psychiatr. Treat. **10**(5), 331–337 (2004)
52. Tversky, A., Kahneman, D.: Judgment under uncertainty: heuristics and biases. Science **185**(4157), 1124–1131 (1974)
53. Wang, S., Zhang, C., Kröse, B., van Hoof, H.: Optimizing adaptive notifications in mobile health interventions systems: reinforcement learning from a data-driven behavioral simulator. J. Med. Syst. **45**(12) (2021)
54. Warburton, D.E., Jamnik, V.K., Bredin, S.S., Gledhill, N.: The physical activity readiness questionnaire for everyone (par-q+) and electronic physical activity readiness medical examination (eparmed-x+). Health Fitness J. Canada **4**(2), 3–17 (2011). https://doi.org/10.14288/hfjc.v4i2.103
55. Wattanapisit, A., Thanamee, S.: Evidence behind 10,000 steps walking. J. Health Res. **31**(3), 241–248 (2017)
56. White, D.K., et al.: The association of obesity with walking independent of knee pain: the multicenter osteoarthritis study. J. Obes. **2012** (2012). https://doi.org/10.1155/2012/261974
57. Wilson, K.E., Dishman, R.K.: Personality and physical activity: a systematic review and meta-analysis. Personality Individ. Differ. **72**, 230–242 (2015)
58. Xu, Y., Zhang, J., Chi, R., Deng, G.: Enhancing customer satisfaction with chatbots: the influence of anthropomorphic communication styles and anthropomorphised roles. Nankai Bus. Rev. Int. **14**(2), 249–271 (2023)
59. Zhou, M., et al.: Personalizing mobile fitness apps using reinforcement learning. In: CEUR Workshop Proceedings, vol. 2068. NIH Public Access (2018)