Promoting metacognitive learning through systematic reflection

Frederic Becker
Max Planck Institute for Intelligent Systems
Tübingen
frederic.becker@tuebingen.mpg.de

Falk Lieder
Max Planck Institute for Intelligent Systems
Tübingen
falk.lieder@tuebingen.mpg.de

Abstract

People are able to learn clever cognitive strategies through trial and error from small amounts of experience. This is facilitated by people's ability to reflect on their own thinking which is known as metacognition. To examine the effects of deliberate systematic metacognitive reflection on how people learn how to plan, the experimental group was guided to systematically reflect on their decision-making process after every third decision. We found that participants assisted by reflection prompts learned to plan better faster. Moreover, we found that reflection led to immediate improvements in the participants' planning strategies. Our preliminary results do suggest that deliberate metacognitive reflection can help people discover clever cognitive strategies from very small amounts of experience. Understanding the role of reflection in human learning is a promising approach for making reinforcement learning more sample efficient in both humans and machines.

1 Introduction

One of the most impressive feats of human intelligence is people's ability to discover and continuously refine their own algorithms [23, 20]. That is, people are able to learn clever cognitive strategies through trial and error. This ability is known as *metacognitive reinforcement learning* [17, 15, 14, 12]. Although work in artificial intelligence has begun to recreate this ability in machines [4], those algorithms still require much more experience to discover good algorithms than people do.

People's superior ability to discover clever cognitive strategies from small amounts of experience appears to be at least partly due to their capacity to reflect on their own thinking which is known as metacognition [22]. Metacognition comprises the monitoring of reasoning, reasoning about reasoning (metareasoning), and the control of reasoning [6, 7, 21] and is an active research topic in both artificial intelligence and psychology [11, 10]. The fact that we can direct people to use different metacognitive strategies by simply asking them a series of questions [8] makes the human mind a convenient platform for research on the potential benefits of different forms of metareasoning.

Although people have the capacity for metacognition, they do not always employ it. Furthermore, even when people engage in metacognition, what they learn depends on which aspects of their thinking they reflect on and how they reflect on it [8]. This raises the question which metacognitive strategies are most conducive to the discovery of clever cognitive strategies. Previous work has developed computational models of how people discover cognitive strategies [22, 15, 14, 20, 12]. These models have given rise to cognitively-inspired reinforcement learning algorithms for automatic algorithm discovery [4]. This approach is known as *metacognitive reinforcement learning*. Despite this progress in reverse-engineering how people discover cognitive strategies, the potential contributions of deliberate metacognitive reflection have yet to be investigated.

35th Conference on Neural Information Processing Systems (NeurIPS 2021), Sydney, Australia.

We investigate this question in the context of reasoning about what to do (i.e., decision-making). Concretely, we investigate how systematic metacognitive reflection affects people's capacity to discover adaptive decision strategies for decisions that require planning multiple steps ahead. As a first step toward investigating which metareasoning algorithms might be most effective, we devised a series of questions that guide people to monitor, reflect on, and control their decision strategies and measured how the metareasoning algorithm entailed by answering those questions affects people's capacity to discover good planning strategies through trial and error, which is a form of metacognitive reinforcement learning [15, 14, 12].

We found that this metacognitive reflection helped participants discover better planning strategies more quickly. This suggests that the metacognitive strategy we tested in our experiment might be a promising approach to make reinforcement learning more sample efficient.

2 Method

The experiment was conducted in accordance to study protocol 429/2021BO2 approved by the Independent Ethics Commission of the Medical Faculty of the University Tübingen.

2.1 Participants

We recruited 41 participants from Amazon Mechanical Turk, 2 (4.9%) of whom were excluded because they required more than three quiz attempts. Each participant was randomly assigned to be either in the control condition or in the experimental condition. This resulted in 19 people in the control condition and 20 in the experimental condition. The control condition spent on average 27 minutes on the experiment whereas the experimental condition spent 34 minutes on average. On average a participant was 35 years old (range: 23-67 years; 18 female). Each participant received a performance dependent bonus of up to $2.5 (mean bonus: $1.5) in addition to a payment of $2.

2.2 Procedure

After participants gave informed consent and completed the need for cognition scale questionnaire [9] and the cognitive reflection test [3], the experiment started with instructions on the experimental task. Their understanding of the instructions was tested via a quiz comprising four basic comprehension questions. If a participant answered one or more questions incorrectly they had to reread the instructions and retake the quiz until they got all answers right. Participants were then informed about the bonus scheme and played 30 trials of the decision-making task described below. After every third trial, except for the last trial, participants were prompted to reflect on their decision (experimental condition) or something irrelevant to the experiment (control condition).

2.3 Materials

2.3.1 Decision-Making Task

Since it is not possible to observe human planning directly, the underlying cognitive process has to be inferred from people's behavior. To this end, we employed the Mouselab Markov Decision Process (Mouselab-MDP) paradigm [5]. The Mouselab-MDP paradigm uses a spatial planning task in which people's information-gathering behavior is highly informative about their planning strategy. In our version of this task, participants were tasked to move a spider from a starting node to one of 18 target nodes. Each possible path consisted of 5 nodes, which contain rewards whose values are initially occluded (see Figure 1). Participants can reveal the value of a reward for a fee of $1 by clicking on the corresponding node. The fact that performing a planning operation requires the participant to first economically acquire the necessary data through clicking makes the resulting data on which nodes a participant clicked on in which order highly informative about what kind of planning strategy they are using (e.g., depth-first search or breadth-first search). Once the spider is moved, clicking is no longer possible. The spider uncovers and collects every reward on its way. The participants' task is to maximize their game score, which is the sum of the rewards collected by the spider minus the amount spent to uncover the values of the rewards. Rewards are drawn from a Gaussian distribution with mean 0 and standard deviation 1, 2, 4, 8, or 32 for nodes that are 1, 2, 3, 4, and 5 steps away from the start node, respectively. This means that the rewards near the start node vary less than the rewards at the target nodes, making it advantageous to start planning at the target nodes. Participants have to solve several different instances of this planning task in a row – each time with a different set of hidden rewards.

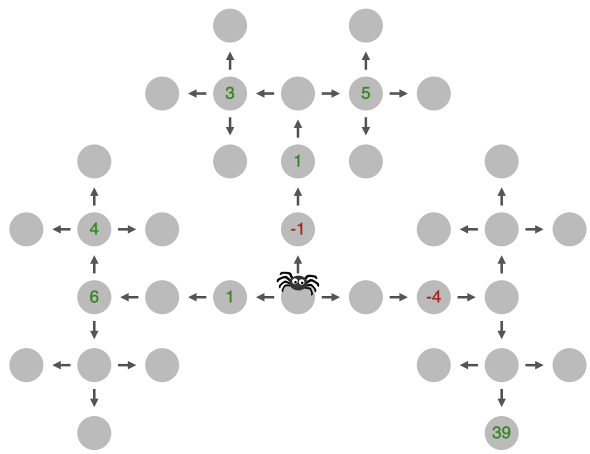

Figure 1: The Mouselab-MDP paradigm. Participants can reveal rewards for a fee to plan a path to one of the end nodes.

2.3.2 Prompts

After every third trial, participants were prompted with a short intervention. In the experimental group, this intervention was designed to foster metacognitive reflection on the planning strategy used in the last 3 trials. It started with the following information: "Please reflect about your planning success in the last three rounds by answering a couple of questions". Participants then received a brief objective feedback on their current performance, displayed for 4 seconds: "Your average score in the last three rounds was X. Your average score in the three rounds before that was Y.". Following Wolfbauer et al. [24], they were guided through different levels of reflection. First, they were asked to *describe* their planning strategy: "How many clicks did you make to decide what to do?", "Where did you click in order to decide what to do?" and "Why did you use that approach?". Second, they were asked to *judge* their planning strategy: "How well do you think your current strategy is working?". Third, they were asked to *evaluate* and *learn* from their previous approach: "Based on what you have learned in the last rounds, what tip could you give to a person who performs this task for the first time?". Fourth, they were asked to *plan* their next approach: "Based on the previous questions, how many clicks do you plan to do in the next rounds?" and "Based on the previous questions, where do you plan to click in the next rounds?". Each question was asked consecutively on its own and each answer had to consist of at least one character, other than that there was no restriction. After this, participants proceeded to the next trial of the decision-making task.

In the control condition, the intervention was designed to have the same duration and elicit similar kinds of activities as the intervention in the experimental condition. It started with the same brief objective feedback of their current performance that was shown to the experimental group. But unlike in the experimental condition, participants in the control condition then answered a question about their personal preferences, such as "What is your favorite food and why?" and then proceeded to the next trial of the decision-making task. On average, the experimental condition answered the entire prompt with 143 characters and took 80 seconds while the control condition answered with 128 characters on average and took 48 seconds.

2.3.3 Questionnaires

We hypothesized that both the attitude towards thinking and the ability to reflect may influence participant performance and the impact of reflection prompts. Therefore, we measured both qualities and included them as covariates in the analysis.

We assessed the degree of satisfaction a participant associates with thinking using the need for cognition (NFC) scale [3]. It consists of 18 items (e.g. "I prefer complex to simple problems") and asks for the degree of approval on a 5-point scale ranging from "extremely uncharacteristic of me" to "extremely characteristic of me". In addition, we included the cognitive reflection test (CRT) [9] to measure participants' "ability or disposition to reflect on a question and resist reporting the first response that comes to mind" ([9], p.1). The CRT consists of 6 questions such as "A bat and a ball cost $1.10 in total. The bat costs $1.00 more than the ball. How much does the ball cost?". Each participant's performance is scored by their number of correct answers.

2.3.4 Computational Microscope

The computational microscope [13] is a computational method for inferring the planning strategies participants use in the Mouselab-MDP paradigm from their information-gathering behavior. It infers a participant's planning strategies by calculating the likelihood for each of 89 predefined planning strategies based on the sequence of clicks performed and a prior that takes the participant's previous strategy choice into account. The predefined planning strategies include, for example, the optimal strategy for this task which starts by exploring the final outcomes and terminates clicking upon finding the maximum value of the reward distribution, breadth-first search, depth-first search, random planning, and no planning. In addition, [13] grouped these planning strategies into 13 different types, including myopic planning strategies (which focus on the reward for the first move), goal-setting strategies (which focus on the rewards at the potential final destinations), and frugal planning strategies (which plan very little). The computational microscope has been empirically validated on the Mouselab-MDP paradigm; it made accurate inferences and was able to detect the effects of feedback on metacognitive learning [13]. We employ the computational microscope to identify possible effects of systematic reflection on the temporal evolution of people's planning strategies.

2.4 Data analysis

To analyse how the observed behavior depends on different factors, we employed linear mixed models (LMM) for continuous data and generalized LMMs (GLMMs) for binary data.

To answer the question of how systematic metacognitive reflection affects individuals' planning performance, we analyzed participants' game score per trial. The employed LMMs regressed the participants' trial-by-trial game scores onto the trial number and the experimental condition. In addition the models included the participants' NFC and CRT scores as covariate and their interactions with the independent variables (see Appendix Table 1). To answer the question of how systematic metacognitive reflection affects individuals' ability to discover adaptive planning strategies, we additionally analyzed the temporal evolution of their planning strategy as measured by the computational microscope, the type of their planning strategy, and its adaptiveness.

The first and most fine-grained descriptor of a participant's planning on a given trial was their *planning strategy*. Using the computational microscope, we inferred which planning strategy a participant used in each trial. To measure the adaptiveness of people's planning strategies, we estimated each strategy's expected game score by its average performance across 100,000 simulations. In a next step, we calculated whether participants changed their strategy between trials, whether such a change was advantageous or disadvantageous in terms of expected game score, and the magnitude of such a change in terms of expected game score. For each of the three variables, we analyzed how they depend on the experimental condition, trial number, NFC score, CRT score, reflection prompt and control prompt (see Appendix Tables 7-9). The latter two described whether an intervention occurred in the trial transition.

The second descriptor was the *type of the planning strategy*. The computational microscope describes 13 different types of planning strategies based on similarity. For each trial, it assigns the planning strategy to a super-ordinate strategy type. We analyzed how the application of a strategy type depends on the experimental condition, trial number and their interactions (see Appendix Table 5). As we employed this procedure for each strategy type, we corrected the p-values with the Benjamini-Hochberg procedure [2]. In addition, we analyzed whether participants switched the type of planning strategy between trials (see Appendix Table 6).

The third descriptor was the *adaptiveness* of the planning strategy. Our sample used 44 different planning strategies in total. Within these, we identified two clusters by applying k-means clustering to the expected game scores of the strategies. In choosing the number of clusters, we followed the average silhouette method [19]. Cluster 1 included 35 strategies whose mean expected score was 33.35; we will refer to these strategies as adaptive planning strategies. Cluster 2 included 9 strategies whose mean expected game score was -1.52; we therefore refer to these strategies as the maladaptive planning strategies. We then analyzed how the application of adaptive planning strategies depend on condition, trial number, NFC score, CRT score and their interactions (see Appendix Table 2). Moreover, we analysed whether participants switched between adaptive and maladaptive planning strategies between trials and if so in which direction using the same model as for strategy and type changes (see Appendix Tables 3 and 4).

The data-analysis was performed in R [18]. We used the packages lme4 [1] for fitting and lmerTest [16] to obtain corresponding p-values. We consider variables with a p-value lower than 0.05 as significant predictors in our models. All details of these analyses and their results are available in the Appendix.

3 Results

3.1 Participants who were guided to reflect on their decisions learned to plan better faster

The experimental group performed better than the control condition ($t(37) = -1.34, p = .186$). The mean game score per trial was 21.6 (Median: 26.5, SD: 32.7) in the control condition and 27.8 (Median: 30, SD: 30.3) in the experimental condition. The usage of adaptive planning strategies (regression coefficient $\beta = 25.68, p < .001$) and the game score ($\beta = 8.39, p < .001$) increased significantly throughout the experiment. The experimental group appeared to learn more rapidly in the beginning and reach a plateau earlier (see Figure 2). To test this assumption we analysed the trials in which the reflection group showed ongoing learning separately. We estimated this phase to be the first 12 trials, since no participant in the experimental group switched to an adaptive planning strategy or vice versa thereafter.

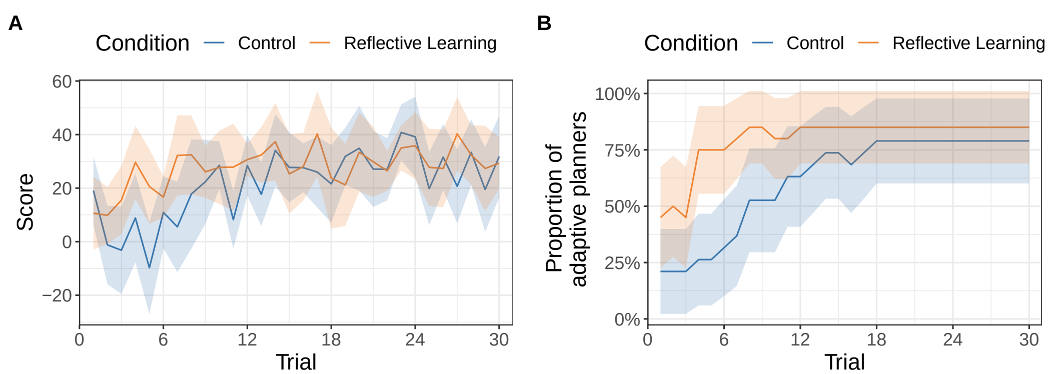

Figure 2: A: Performance curve: Mean game score per trial and condition. B: Learning curve: Proportion of participants planning with an adaptive planning strategy per trial and condition. The vertical lines represent the occurrence of reflection or control prompts after every 3rd trial. The dashed areas convey 95% confidence intervals.

In the first 12 trials, the experimental condition performed significantly better than the control condition ($t(37) = -2.80, p = .008$). The mean game score per trial was 11.3 (Median: 14, SD: 34.4) in the control condition and 23.3 (Median: 25.5, SD: 29.3) in the experimental condition (see Figure 3A). Reflection prompts significantly increased the proportion of people who adopted adaptive planning strategies ($\beta = 6.87, p = .029$, see Figure 3B). The control group applied an adaptive strategy in 39% of trials, whereas the experimental group did so in 72% of trials. Again we found a positive effect of trial number on game score and the use of adaptive planning strategies (see Appendix Table 11).

The increased use of adaptive planning strategies in the experimental condition might be due to increased switching from maladaptive to adaptive planning strategies. Within the first six trials, 15.7% of participants in the control condition performed one or more switches, whereas 50% of

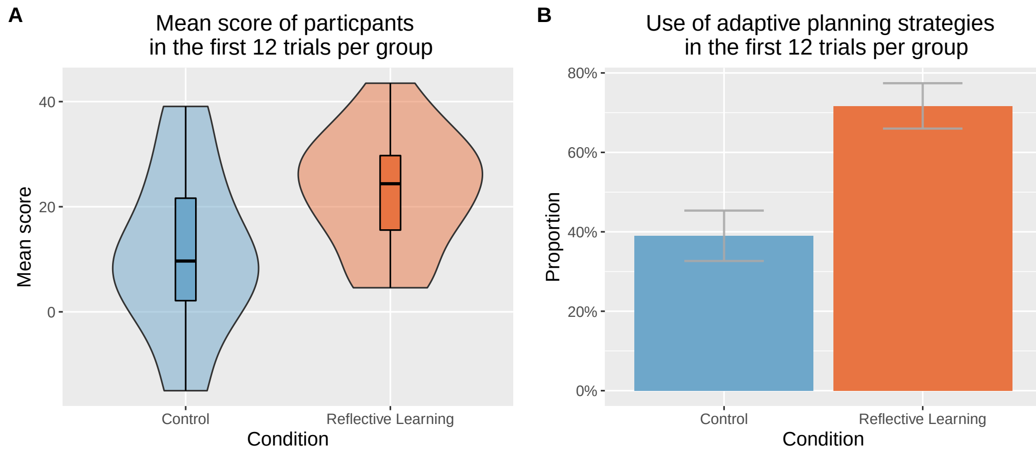

Figure 3: Systematic metacognitive reflection helped people to learn better faster A: Performance: Mean game score in the first 12 trials per condition presented as box-plot. B: Adaptiveness: Use of adaptive planning strategies in the first 12 trials per group. The error bars convey 95% confidence intervals.

participants in the experimental condition did so. There was a positive (and almost statistically significant) effect of reflection on the frequency of switches in the first 12 trials ($\beta = 3.03, p = .052$) as well as a significant negative effect of the interaction of the experimental condition and trial number ($\beta = -0.61, p = .021$). This suggests that the experimental condition performed exploration earlier and more frequently than the control condition. This also implies that in the beginning of the experiment the experimental group learned more than the control group. We did not find this pattern for switches between more fine-grained planning strategies or types of planning strategies (see Appendix Table 16 and 15).

Across all 30 trials, we found that the experimental group performed significantly more switches from an adaptive to a maladaptive planning strategy ($\beta = -6.69, p = .044$) than the control group. Curiously, the experimental group also made more switches from a strategy with a higher expected game score to one with lower expected game score ($\beta = -2.5, p = .041$) than the control group. This effect did not remain significant when we only considered the first 12 trials ($\beta = -5.14, p = .098$) and when we analysed continuous decreases of expected game score due to strategy switches ($\beta = -15.45, p = .066$). This suggest that in later trials, where adaptive strategies were already found, increased exploration led to less advantageous switching.

3.2 Reflection prompts led to immediate beneficial changes in planning on the next trial

Since the groups differed in terms of their performance and their use of adaptation strategies, we now examine the immediate effects of reflection prompts. Considering all 30 trials as well as only the first 12 trials, we found that the presentation of a reflection prompt increased the probability of an immediate change in the strategy, the strategy type, or the adaptiveness of the strategy from the previous trial to the next trial (all $\beta : 1.46 - 3.35$, all $p \leq .041$). This was not the case for control prompts (all $\beta : -0.5 - 3.05$, all $p \geq .148$). Hence, the effect of the reflection prompts cannot be solely due to the presentation of performance feedback or the interruption of the task, because these elements were also present in the control prompt. Instead, this difference is most likely due to the metacognitive reflection script.

In addition, the difference in expected game score when switching between strategies was more positive when the switch followed a reflection prompt (trials 1-12 trials: $\beta : 47.13, p = .008$, trials 1-30: $\beta : 22.08, p = .006$). In the first 12 trials, the average gain in expected game score when performing a switch after a reflection prompt was 19.74 (Median: 18.4, SD: 20.2) and 3.9 (Median: 0.2, SD: 22.4) in the remaining trials of the experimental condition. Reflection prompts thus elicited immediate and highly beneficial changes in planning strategies. This further corroborates the interpretation that the higher performance of the experimental condition was due to the reflection prompts rather than preexisting differences between the two groups.

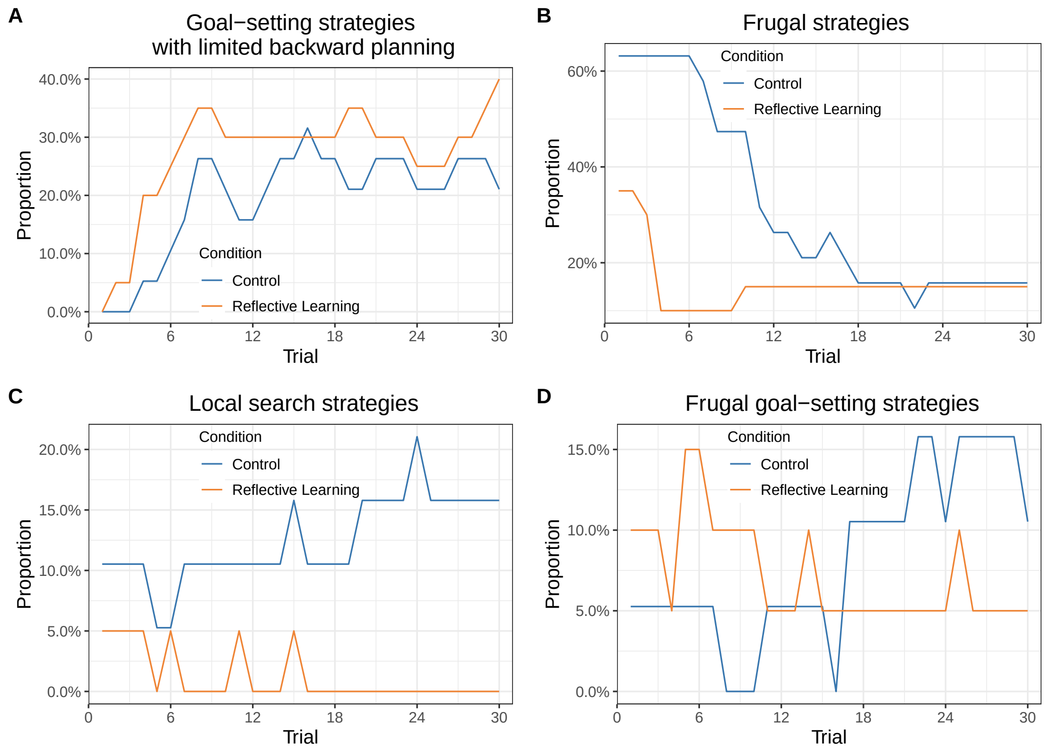

Figure 4: Application of different types of planning strategies over time per condition: A: Goal-setting strategies with limited backward planning, start to explore final outcomes and terminate depending whether a high enough reward has been found. B: Frugal planning strategies, explore very little or not at all. C: Local search strategies, focus on exploring single branches. D: Frugal goal-setting planning strategies, start to explore final outcomes but explore very little overall.

3.3 How did reflection affect planning ?

Having seen *that* reflection prompts improved human planning immediately, we now inspect *how* this improvement was achieved. Identifying how reflection changes people's planning strategies is an important first step toward understanding the mechanisms of reflective learning and might point to opportunities to improve our reflection script. Due to group differences in the initial strategy use and the low sample size, we did not have enough statistical power for systematic hypothesis testing. So we only report descriptive statistics and exploratory analyses instead.

We first compared the experimental condition and the control condition in terms of their most frequently used strategies. The most frequently used strategy type of the experimental group was goal-setting strategies with limited backward planning. These highly adaptive strategies first inspect the potential final destinations to select one or more potential goals and may then inspect some of the nodes along the path to the chosen goal(s). The experimental group used this adaptive strategy type in 27.2% of the trials whereas the control group used it in only 19.5% of the trials (see Figure 4A). The groups differed the most in how often they used frugal planning strategies, that is strategies that perform very little planning or no planning at all (see Figure 4B). Frugal planning was the most frequent strategy type of the control group, which used it in 31.8% of all trials, but the experimental group used it less than half as often, which used it in only 15.8% of all trials.

Six out of eight strategy types became more frequent over time or less frequent over time (see Appendix Table 5). For two of them, the changes in the control group and the experimental group pointed in opposite directions (see Figure 4C and D). Local-search strategies focus on exploring branches for which most rewards are already known. In the experimental condition, participants' use of such strategies decreased over time ($\beta = -2.19, p = .005$). By contrast, in the control condition, participants' use of such strategies increased over time ($\beta = 0.41, p = .019$), see Figure 4C. As shown in Figure 4D, we found the same pattern for frugal goal-setting strategies (control group:

$\beta = 1.04, p < .001$, experimental group: $\beta = -2.3, p < .001$). These strategies are similar to the goal-setting strategies with limited backward planning in that they focus on the rewards of potential final destinations; but they perform less planning overall.

The differences in how often the two groups relied on frugal strategies and frugal goal-setting strategies suggests that reflection might led people to plan more. This hypothesis still needs to be tested in a larger experiment; with a larger data set, analyses of how fast certain types of strategies were adopted or abandoned may lead to additional insights.

3.4 The influence of Need for Cognition (NFC) and Cognitive Reflection (CRT)

Looking at all 30 trials, we found that high NFC reduced the rate at which participants' game score increased with practice ($\beta = -2.64, p = .036$) and the rate at which they adopted adaptive strategies ($\beta = -12.21, p < .001$). Interestingly, the opposite was the case for the experimental condition (game score: $\beta = 4.96, p = .004$, adaptiveness: $\beta = 12.58, p =< .001$). This suggests that, on the one hand, high NFC made learning more difficult in the control condition but, on the other hand, made reflection prompts more effective in the experimental condition. This might be because a higher need for cognition makes the experimental condition more engaging and the control condition less engaging.

In addition, we found that in the control condition people who performed better on the CRT adopted adaptive strategies faster than people with a lower CRT score ($\beta = 15.26, p < .001$). But when reflection prompts were provided then people with low CRT scores learned just as fast as people with high CRT scores ($\beta = -13.17, p < .001$). This suggests that the reflection prompts were especially beneficial for people with low CRT scores. We also found a positive correlation between participants' performance on the CRT and the number of clicks they performed prior to the first intervention ($r(115) = .24, p = .009$.), hence people with low CRT tend to underplan in the beginning. We thus assume that reflection prompts were especially beneficial for people who tend to make less-deliberate decisions.

4 Discussion

In summary, we found that a specific method for deliberate metacognitive reflection boosted our participants' (metacognitive) reinforcement learning to a significant extent. The main limitation of our study is its low sample size. As a consequence, the two groups differed in their initial performance and initial strategy use. This complicates the interpretation of our findings. Our conclusions should therefore be taken with a grain of salt. Although more research is needed to assess the robustness and reproducibility of our findings, our preliminary results do suggest that deliberate metacognitive reflection can help people discover clever cognitive strategies from very small amounts of experience. This might also be true of other forms of reinforcement learning. More generally, the high sample efficiency of human learning could be partly due to sophisticated learning mechanisms that rely on reasoning. This makes investigating the role of (metacognitive) reflection in human (metacognitive) learning a promising starting point for making reinforcement learning more sample efficient in both humans and machines.

In future work, we will measure the degree to which participants engage with the reflection prompts and investigate whether it moderates the effect of reflection on learning. In addition, we plan to include a second passive control condition in which participants work on the task continuously without breaks to rule out potential negative disruption effects. Moreover, future work will investigate how the benefits of metacognitive reflection depend on what exactly people reflect on and how they reflect on it. We are planning to investigate this question in the Mouselab-MDP paradigm by varying the reflection script across different experimental conditions. This will hopefully provide AI researchers with some inspiration as to which kinds of metareasoning might be most worthwhile to incorporate into intelligent systems. Developing computational models of metacognitive reflection and its contributions to metacognitive learning is an exciting direction for future research. In the long run, this line of research lead to metacognitive intelligent systems that can discover and continuously improve their own algorithms.

References

[1] D. Bates, M. Mächler, B. Bolker, and S. Walker. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1):1–48, 2015.

[2] Y. Benjamini and Y. Hochberg. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, 57(1):289–300, 1995.

[3] J. T. Cacioppo, R. E. Petty, and C. F. Kao. The efficient assessment of need for cognition. *Journal of Personality Assessment*, 48(3):306–307, 1984. PMID: 16367530.

[4] F. Callaway, S. Gul, P. Krueger, T. L. Griffiths, and F. Lieder. Learning to select computations. In *Uncertainty in Artificial Intelligence: Proceedings of the Thirty-Fourth Conference*, 2018.

[5] F. Callaway, F. Lieder, P. M. Krueger, and T. L. Griffiths. Mouselab-MDP: A new paradigm for tracing how people plan. In *The 3rd Multidisciplinary Conference on Reinforcement Learning and Decision Making, Ann Arbor, MI*, 2017.

[6] M. Cox and A. Raja. Metareasoning: A manifesto. *BBN Technical*, 2007.

[7] M. T. Cox and A. Raja. *Metareasoning: Thinking about thinking*. MIT Press, 2011.

[8] S. Ellis, B. Carette, F. Anseel, and F. Lievens. Systematic reflection: Implications for learning from failures and successes. *Current Directions in Psychological Science*, 23(1):67–72, 2014.

[9] S. Frederick. Cognitive reflection and decision making. *Journal of Economic Perspectives*, 19(4):25–42, December 2005.

[10] S. J. Gershman, E. J. Horvitz, and J. B. Tenenbaum. Computational rationality: A converging paradigm for intelligence in brains, minds, and machines. *Science*, 349(6245):273–278, 2015.

[11] T. L. Griffiths, F. Callaway, M. B. Chang, E. Grant, P. M. Krueger, and F. Lieder. Doing more with less: meta-reasoning and meta-learning in humans and machines. *Current Opinion in Behavioral Sciences*, 29:24–30, 2019.

[12] R. He, Y. R. Jain, and F. Lieder. Measuring and modelling how people learn how to plan and how people adapt their planning strategies the to structure of the environment. In *International Conference on Cognitive Modeling*, 2021.

[13] Y. R. Jain, F. Callaway, T. L. Griffiths, P. Dayan, P. M. Krueger, and F. Lieder. A computational process-tracing method for measuring people's planning strategies and how they change over time. *Preprint, DOI: 10.13140/RG.2.2.25963.05925/1*, 2021.

[14] Y. R. Jain, S. Gupta, V. Rakesh, P. Dayan, F. Callaway, and F. Lieder. How do people learn how to plan? In *Conference on cognitive computational neuroscience (ccn 2019)*, pages 826–829, 2019.

[15] P. M. Krueger, F. Lieder, and T. Griffiths. Enhancing metacognitive reinforcement learning using reward structures and feedback. In *CogSci*, 2017.

[16] A. Kuznetsova, P. B. Brockhoff, and R. H. B. Christensen. lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, 82(13):1–26, 2017.

[17] F. Lieder and T. L. Griffiths. Strategy selection as rational metareasoning. *Psychological Review*, 124(6):762–794, 2017.

[18] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2017.

[19] P. J. Rousseeuw. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics*, 20:53–65, 1987.

[20] J. S. Rule, J. B. Tenenbaum, and S. T. Piantadosi. The child as hacker. *Trends in cognitive sciences*, 2020.

[21] S. Russell and E. Wefald. Principles of metareasoning. *Artificial intelligence*, 49(1-3):361–395, 1991.

[22] J. Shrager and R. S. Siegler. Scads: A model of children's strategy choices and strategy discoveries. *Psychological science*, 9(5):405–410, 1998.

[23] R. S. Siegler. Strategic development. *Trends in Cognitive Sciences*, 3(11):430–435, 1999.

[24] I. Wolfbauer, V. Pammer-Schindler, and C. P. Rosé. Rebo junior: Analysis of dialogue structure quality for a reflection guidance chatbot. In *EC-TEL (Impact Papers)*, 2020.

Checklist

1. For all authors...
   (a) Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope? [Yes] The abstract adequately describes our contributions elaborated in the main text.
   (b) Did you describe the limitations of your work? [Yes] See section 4
   (c) Did you discuss any potential negative societal impacts of your work? [No] We cannot think of any possible negative impact that our research might have on society.
   (d) Have you read the ethics review guidelines and ensured that your paper conforms to them? [Yes]
2. If you are including theoretical results...
   (a) Did you state the full set of assumptions of all theoretical results? [N/A]
   (b) Did you include complete proofs of all theoretical results? [N/A]
3. If you ran experiments...
   (a) Did you include the code, data, and instructions needed to reproduce the main experimental results (either in the supplemental material or as a URL)? [N/A]
   (b) Did you specify all the training details (e.g., data splits, hyperparameters, how they were chosen)? [N/A]
   (c) Did you report error bars (e.g., with respect to the random seed after running experiments multiple times)? [N/A]
   (d) Did you include the total amount of compute and the type of resources used (e.g., type of GPUs, internal cluster, or cloud provider)? [N/A]
4. If you are using existing assets (e.g., code, data, models) or curating/releasing new assets...
   (a) If your work uses existing assets, did you cite the creators? [Yes] See section 2.3.3 and 2.3.4
   (b) Did you mention the license of the assets? [No]
   (c) Did you include any new assets either in the supplemental material or as a URL? [No]
   (d) Did you discuss whether and how consent was obtained from people whose data you're using/curating? [Yes] See section 2.2
   (e) Did you discuss whether the data you are using/curating contains personally identifiable information or offensive content? [No] The data was fully anonymized.
5. If you used crowdsourcing or conducted research with human subjects...
   (a) Did you include the full text of instructions given to participants and screenshots, if applicable? [Yes] Partly, see section 2.3.2
   (b) Did you describe any potential participant risks, with links to Institutional Review Board (IRB) approvals, if applicable? [Yes] See 2
   (c) Did you include the estimated hourly wage paid to participants and the total amount spent on participant compensation? [Yes] We reported the average duration and the average compensation, see section 2.1

Appendix

Table 1: Linear mixed-model for game score (30 trials)

Fixed Effect	Estimate	95% CI	df	t	p	Significance
Intercept	22.65	[15.9, 29.4]	35	6.39	<.001	*
Condition	4.11	[-5.83, 14.05]	35	0.79	.436	
Trial	8.39	[5.74, 11.04]	1125	6.2	<.001	*
CRT	2.52	[-2.41, 7.46]	35	0.97	.337	
NFC	-0.31	[-4.95, 4.34]	35	-0.12	.901	
Condition:Trial	-3.99	[-7.59, -0.39]	1125	-2.17	.03	*
CRT:Trial	-1.48	[-4.37, 1.4]	1125	-1.01	.315	
NFC:Trial	-2.64	[-5.11, -0.18]	1125	-2.1	.036	*
Condition:CRT:Trial	1.84	[-1.82, 5.5]	1125	0.99	.325	
Condition:NFC:Trial	4.96	[1.6, 8.31]	1125	2.89	.004	*

Table 2: Generalized linear mixed-model for the use of adaptive strategies (30 trials)

Fixed Effect	Estimate	95% CI	z	p	Significance
Intercept	13.66	[3.15, 28.46]	2.4	.016	*
Condition	0.69	[-12.51, 12.8]	0.12	.906	
Trial	25.68	[17.56, 37.78]	5.19	<.001	*
CRT	3.09	[-1.89, 9.89]	1.36	.174	
NFC	-1.56	[-8.87, 3.71]	-0.53	.598	
Condition:Trial	-23.26	[-35.26, -15.07]	-4.71	<.001	*
CRT:Trial	15.26	[10.16, 22.66]	4.99	<.001	*
NFC:Trial	-12.21	[-17.92, -8.28]	-5.18	<.001	*
Condition:CRT:Trial	-13.17	[-20.58, -8]	-4.27	<.001	*
Condition:NFC:Trial	12.58	[8.54, 18.36]	5.23	<.001	*

Table 3: Generalized linear mixed-model for changes between adaptive and maladaptive strategies (30 trials)

Fixed Effect	Estimate	95% CI	z	p	Significance
Intercept	-2.78	[-3.91, -1.93]	-5.13	<.001	*
Condition	0.94	[-0.92, 2.71]	0.99	.321	
Trial	-0.05	[-0.14, 0.02]	-1.34	.181	
CRT	-0.47	[-1.27, 0.29]	-1.18	.238	
NFC	-0.18	[-0.79, 0.45]	-0.57	.572	
ReflectionPrompt	2.9	[1.02, 5.02]	2.87	.004	*
ControlPrompt	0.24	[-1.94, 2.14]	0.23	.815	
Condition:Trial	-0.36	[-0.78, -0.1]	-1.95	.051	
NFC:Trial	0.02	[-0.04, 0.07]	0.54	.588	
CRT:Trial	0.02	[-0.06, 0.08]	0.51	.608	
ReflectionPrompt:Trial	-0.25	[-0.7, 0.04]	-1.39	.165	
ControlPrompt:Trial	-0.06	[-0.26, 0.08]	-0.66	.508	
Condition:NFC:Trial	0.01	[-0.1, 0.15]	0.22	.826	
Condition:CRT:Trial	0.17	[0.01, 0.39]	1.74	.082	

Table 4: Generalized linear mixed-model for changes towards an adaptive strategy (30 trials)

Fixed Effect	Estimate	95% CI	z	p	Significance
Intercept	9.09	[2.68, 26.79]	2.01	.045	*
Condition	-6.69	[-15.38, -1.51]	-2.01	.044	*
Trial	-0.5	[-1.34, -0.05]	-1.63	.103	
CRT	1.99	[0.3, 4.72]	1.92	.055	
NFC	-0.43	[-1.73, 0.68]	-0.74	.46	
ReflectionPrompt	1.16	[-2.05, 5.22]	0.68	.497	
ControlPrompt	0.18	[-3.61, 4.47]	0.09	.927	

Table 5: Generalized linear mixed-model for types of planning strategies (30 trials)

Fixed Effect	Estimate	95% CI	z	p (corrected)	Significance
Goal-setting with limited backward planning					
Intercept	-7.05	[-12.45, -3.35]	-2.76	.008	*
Condition	1.91	[-2.7, 6.23]	0.89	.497	
Trial	1.17	[0.77, 1.64]	5.35	<.001	*
Condition:Trial	-0.37	[-0.93, 0.17]	-1.32	.248	
Local search strategies					
Intercept	-4.69	[-9.22, -2.8]	-3.51	<.001	*
Condition	-3.84	[-7.28, -0.64]	-2.5	.098	
Trial	0.41	[0.09, 0.76]	2.45	.019	*
Condition:Trial	-2.19	[-3.94, -0.98]	-3	.005	*
Other miscellaneous strategies					
Intercept	-7.48	[-12.52, -3.88]	-3.44	<.001	*
Condition	2.21	[-2.23, 6.3]	1.08	.446	
Trial	0.93	[0.56, 1.34]	4.75	<.001	*
Condition:Trial	-0.08	[-0.58, 0.41]	-0.31	.755	
Forward planning strategies					
Intercept	-34.7	[-94.77, -18.34]	-2.63	.01	*
Condition	22.05	[6.38, 79.33]	1.79	.217	
Trial	-14.9	[-49.4, -5.91]	-2.01	.05	
Condition:Trial	14.02	[5.02, 48.5]	1.89	.093	
Frugal goal-setting strategies					
Intercept	-10.29	[-15.74, -6.88]	-5.47	<.001	*
Condition	0.01	[-4.37, 4.67]	0.01	.994	
Trial	1.04	[0.59, 1.55]	4.27	<.001	*
Condition:Trial	-2.3	[-3.33, -1.43]	-4.79	<.001	*
Myopic planning strategies					
Intercept	-8.74	[-13.32, -5.28]	-4.7	<.001	*
Condition	-0.25	[-3.98, 3.62]	-0.14	.994	
Trial	0.07	[-0.37, 0.52]	0.31	.754	
Condition:Trial	-1.84	[-2.69, -1.09]	-4.55	<.001	*
Frugal planning strategies					
Intercept	-3.08	[-6.9, -0.39]	-2.14	.033	*
Condition	-3.6	[-8.02, 0.31]	-1.74	.217	
Trial	-2.4	[-2.96, -1.92]	-9.06	<.001	*
Condition:Trial	1.91	[1.25, 2.6]	5.59	<.001	*
Strategy that explores immediate rewards on the paths to the best final outcomes with satisficing					
Intercept	-10.13	[-16.56, -6.31]	-4.63	<.001	*
Condition	-3.1	[-10.87, 2.67]	-1.08	.446	
Trial	1.39	[0.68, 2.29]	3.45	<.001	*
Condition:Trial	1.65	[-0.58, 5.32]	1.15	.287	

Table 6: Generalized linear mixed-model for changes between types of planning strategies (30 trials)

Fixed Effect	Estimate	95% CI	z	p	Significance
Intercept	-2.54	[-3.43, -1.75]	-6.02	<.001	*
Condition	0.56	[-0.64, 1.77]	0.93	.352	
Trial	-0.01	[-0.06, 0.04]	-0.55	.583	
CRT	-0.37	[-0.9, 0.15]	-1.41	.16	
NFC	-0.04	[-0.53, 0.48]	-0.14	.886	
ReflectionPrompt	1.46	[0.31, 2.62]	2.52	.012	*
ControlPrompt	-0.5	[-2.02, 0.86]	-0.7	.485	
Condition:Trial	-0.06	[-0.14, 0.02]	-1.55	.12	
NFC:Trial	0.03	[0, 0.06]	1.73	.084	
CRT:Trial	0.01	[-0.03, 0.04]	0.36	.715	
ReflectionPrompt:Trial	-0.07	[-0.16, 0.01]	-1.73	.083	
ControlPrompt:Trial	0.04	[-0.04, 0.12]	1.05	.292	
Condition:NFC:Trial	0.02	[-0.02, 0.07]	1.01	.315	
Condition:CRT:Trial	0.04	[0, 0.08]	1.72	.086	

Table 7: Generalized linear mixed-model for changes between planning strategies (30 trials)

Fixed Effect	Estimate	95% CI	z	p	Significance
Intercept	-2.49	[-3.37, -1.7]	-5.99	<.001	*
Condition	0.65	[-0.53, 1.84]	1.1	.273	
Trial	-0.01	[-0.06, 0.04]	-0.38	.701	
CRT	-0.25	[-0.77, 0.25]	-0.99	.324	
NFC	0.05	[-0.44, 0.56]	0.19	.846	
ReflectionPrompt	1.35	[0.24, 2.45]	2.43	.015	*
ControlPrompt	-0.46	[-1.89, 0.85]	-0.67	.5	
Condition:Trial	-0.07	[-0.14, 0.01]	-1.82	.069	
NFC:Trial	0.02	[-0.01, 0.05]	1.36	.175	
CRT:Trial	0	[-0.04, 0.03]	-0.04	.967	
ReflectionPrompt:Trial	-0.06	[-0.15, 0.01]	-1.61	.107	
ControlPrompt:Trial	0.04	[-0.03, 0.12]	1.13	.257	
Condition:NFC:Trial	0.02	[-0.02, 0.07]	0.93	.353	
Condition:CRT:Trial	0.04	[0, 0.09]	2.02	.044	*

Table 8: Generalized linear mixed-model for changes towards a planning strategy with higher expected game score (30 trials)

Fixed Effect	Estimate	95% CI	z	p	Significance
Intercept	2.37	[0.75, 4.17]	2.45	.014	*
Condition	-2.5	[-4.8, -0.49]	-2.05	.041	*
Trial	-0.08	[-0.19, 0.01]	-1.5	.134	
CRT	0.57	[-0.36, 1.56]	1.17	.244	
NFC	-0.32	[-1.46, 0.7]	-0.6	.55	
ReflectionPrompt	1.82	[-0.26, 4.22]	1.61	.107	
ControlPrompt	4.69	[0.07, 12.11]	1.57	.116	
Condition:Trial	0.11	[-0.03, 0.26]	1.46	.144	
NFC:Trial	0.02	[-0.04, 0.09]	0.72	.473	
CRT:Trial	0.01	[-0.07, 0.09]	0.15	.877	
ReflectionPrompt:Trial	-0.06	[-0.22, 0.09]	-0.8	.424	
ControlPrompt:Trial	-0.28	[-0.48, -0.08]	-1.78	.075	
Condition:NFC:Trial	-0.03	[-0.11, 0.04]	-0.87	.385	
Condition:CRT:Trial	-0.05	[-0.13, 0.02]	-1.29	.196	

Table 9: Linear mixed model for the difference in expected game score with changes in planning strategies. (30 trials)

Fixed Effect	Estimate	95% CI	df	t	p	Significance
Intercept	20.03	[8.67, 31.39]	99	3.26	.002	*
Condition	-15.45	[-30.83, -0.08]	99	-1.86	.066	
Trial	-0.65	[-1.35, 0.05]	99	-1.73	.087	
CRT	2.46	[-3.86, 8.77]	99	0.72	.473	
NFC	-8.16	[-14.96, -1.36]	99	-2.22	.029	*
ReflectionPrompt	22.08	[7.63, 36.53]	99	2.83	.006	*
ControlPrompt	15.17	[-6.46, 36.8]	99	1.3	.198	
Condition:Trial	0.47	[-0.54, 1.48]	99	0.86	.39	
NFC:Trial	0.33	[-0.11, 0.78]	99	1.39	.169	
CRT:Trial	0.12	[-0.44, 0.67]	99	0.39	.7	
ReflectionPrompt:Trial	-1.02	[-2.09, 0.04]	99	-1.79	.077	
ControlPrompt:Trial	-0.82	[-2.03, 0.39]	99	-1.26	.211	
Condition:NFC:Trial	0.02	[-0.47, 0.52]	99	0.09	.932	
Condition:CRT:Trial	-0.25	[-0.76, 0.26]	99	-0.89	.374	

Table 10: Linear mixed-model for game score (first 12 trials)

Fixed Effect	Estimate	95% CI	df	t	p	Significance
Intercept	12.98	[6.89, 19.06]	35	4.06	<.001	*
Condition	8.76	[-0.2, 17.71]	35	1.86	.071	
Trial	5.84	[1.38, 10.3]	423	2.56	.011	*
CRT	4.03	[-0.41, 8.48]	35	1.73	.093	
NFC	-0.41	[-4.59, 3.78]	35	-0.19	.854	
Condition:Trial	0.25	[-5.81, 6.32]	423	0.08	.935	
CRT:Trial	-3.69	[-8.56, 1.17]	423	-1.48	.14	
NFC:Trial	-2.83	[-6.99, 1.32]	423	-1.33	.184	
Condition:CRT:Trial	4.35	[-1.82, 10.51]	423	1.38	.17	
Condition:NFC:Trial	4.25	[-1.4, 9.9]	423	1.47	.143	

Table 11: Generalized linear mixed-model for the use of adaptive strategies (first 12 trials)

Fixed Effect	Estimate	95% CI	z	p	Significance
Intercept	-2.23	[-6.67, 1.19]	-1.25	.213	
Condition	6.87	[1.78, 14.52]	2.19	.029	*
Trial	3.52	[2.19, 5.99]	3.93	<.001	*
CRT	-0.32	[-3.04, 2.43]	-0.25	.803	
NFC	0.41	[-2.07, 3.01]	0.35	.727	
Condition:Trial	-2.1	[-4.59, -0.59]	-2.24	.025	*
CRT:Trial	0.2	[-1.05, 2.23]	0.26	.793	
NFC:Trial	-0.34	[-1.59, 0.87]	-0.58	.563	
Condition:CRT:Trial	0.74	[-1.35, 2.2]	0.88	.377	
Condition:NFC:Trial	0.09	[-1.32, 1.48]	0.13	.896	

Table 12: Generalized linear mixed-model for changes between adaptive and maladaptive strategies (first 12 trials)

Fixed Effect	Estimate	95% CI	z	p	Significance
Intercept	-5.31	[-7.47, -3.51]	-4.23	<.001	*
Condition	3.03	[0.21, 5.26]	1.95	.052	
Trial	0.3	[0.05, 0.56]	2.05	.04	*
CRT	-0.82	[-1.91, 0.28]	-1.42	.155	
NFC	-0.19	[-1.12, 0.77]	-0.4	.688	
ReflectionPrompt	3.35	[0.99, 5.81]	2.73	.006	*
ControlPrompt	3.05	[-1.34, 7.23]	1.45	.148	
Condition:Trial	-0.61	[-1.17, -0.19]	-2.31	.021	*
NFC:Trial	-0.02	[-0.16, 0.11]	-0.33	.738	
CRT:Trial	0.04	[-0.14, 0.2]	0.51	.609	
ReflectionPrompt:Trial	-0.36	[-0.87, 0.06]	-1.53	.126	
ControlPrompt:Trial	-0.56	[-1.45, -0.02]	-1.5	.133	
Condition:NFC:Trial	0.06	[-0.08, 0.22]	0.81	.419	
Condition:CRT:Trial	0.23	[0.03, 0.48]	1.99	.047	*

Table 13: Generalized linear mixed-model for changes towards an adaptive strategy (first 12 trials)

Fixed Effect	Estimate	95% CI	z	p	Significance
Intercept	7.53	[-0.86, 15.93]	1.76	.078	
Condition	-5.14	[-11.23, 0.94]	-1.66	.098	
Trial	-0.46	[-1.09, 0.18]	-1.4	.16	
CRT	1.55	[-0.48, 3.58]	1.5	.134	
NFC	-0.23	[-1.39, 0.93]	-0.39	.698	
ReflectionPrompt	0.96	[-2.17, 4.09]	0.6	.547	
ControlPrompt	30.22	[-8851.23, 8881.39]	0	1	

Table 14: Generalized linear mixed-model for types of planning strategies (first 12 trials)

Fixed Effect	Estimate	95% CI	z	p (corrected)	Significance
Goal-setting with limited backward planning					
Intercept	-11.45	[-17.11, -7.65]	-5.51	<.001	*
Condition	1.78	[-2.57, 6.73]	0.84	.491	
Trial	2.57	[1.46, 4.12]	3.88	<.001	*
Condition:Trial	-0.75	[-2.41, 0.63]	-0.99	.45	
Local search strategies					
Intercept	-7.28	[-12.44, -3.69]	-3.43	<.001	*
Condition	-1.44	[-6.02, 2.55]	-0.81	.491	
Trial	0.05	[-0.52, 0.62]	0.17	.868	
Condition:Trial	-0.93	[-2.33, 0.23]	-1.47	.329	
Forward planning strategies					
Intercept	-17.35	[-40.56, -10.24]	-3.09	.002	*
Condition	5.67	[-5.2, 26.73]	1.09	.491	
Trial	-4.62	[-17.38, -1.3]	-1.64	.176	
Condition:Trial	4.76	[1.34, 17.52]	1.68	.326	
Other miscellaneous strategies					
Intercept	-10.76	[-16.17, -6]	-4.45	<.001	*
Condition	3.32	[-0.61, 7.8]	1.63	.357	
Trial	3.08	[1.73, 5.04]	3.73	<.001	*
Condition:Trial	-2.58	[-4.6, -1.13]	-3	.019	*
Frugal goal-setting strategies					
Intercept	-11.03	[-24.26, -7.22]	-4.03	<.001	*
Condition	0.99	[-6.25, 12.61]	0.35	.723	
Trial	-0.52	[-1.61, 0.44]	-1.02	.358	
Condition:Trial	0.1	[-1.15, 1.42]	0.16	.871	
Myopic planning strategies					
Intercept	-8.05	[-13.47, -4.74]	-3.8	<.001	*
Condition	1.64	[-2.15, 5.5]	0.88	.491	
Trial	-0.52	[-1.38, 0.24]	-1.29	.276	
Condition:Trial	-0.33	[-1.3, 0.68]	-0.66	.592	
Frugal planning strategies					
Intercept	-0.2	[-3.98, 3.01]	-0.13	.895	
Condition	-6.37	[-11.67, -2.19]	-2.73	.044	*
Trial	-1.59	[-2.33, -0.98]	-4.69	<.001	*
Condition:Trial	0.5	[-0.4, 1.42]	1.09	.45	

Table 15: Generalized linear mixed-model for changes between types of planning strategies (first 12 trials)

Fixed Effect	Estimate	95% CI	z	p	Significance
Intercept	-3.59	[-5.03, -2.4]	-4.64	<.001	*
Condition	1.53	[-0.29, 3.36]	1.59	.111	
Trial	0.17	[-0.01, 0.36]	1.74	.082	
CRT	-0.45	[-1.24, 0.3]	-1.14	.254	
NFC	-0.08	[-0.8, 0.69]	-0.21	.837	
ReflectionPrompt	2.42	[0.54, 4.33]	2.5	.013	*
ControlPrompt	0.78	[-2.66, 3.54]	0.5	.618	
Condition:Trial	-0.19	[-0.43, 0.06]	-1.47	.143	
NFC:Trial	0	[-Inf, 0.11]	-0.01	.996	
CRT:Trial	0.02	[-0.12, 0.14]	0.25	.8	
ReflectionPrompt:Trial	-0.27	[-0.6, 0.02]	-1.63	.103	
ControlPrompt:Trial	-0.16	[-0.63, 0.3]	-0.69	.489	
Condition:NFC:Trial	0.07	[-0.03, 0.18]	1.36	.174	
Condition:CRT:Trial	0.04	[-0.06, 0.16]	0.75	.454	

Table 16: Generalized linear mixed-model for changes between planning strategies (first 12 trials)

Fixed Effect	Estimate	95% CI	z	p	Significance
Intercept	-3.34	[-4.92, -2.06]	-4.64	<.001	*
Condition	1.35	[-0.43, 3.24]	1.47	.143	
Trial	0.15	[-0.03, 0.34]	1.61	.108	
CRT	-0.44	[-1.21, 0.31]	-1.14	.255	
NFC	-0.15	[-0.87, 0.59]	-0.41	.679	
ReflectionPrompt	1.91	[0.07, 3.77]	2.05	.041	*
ControlPrompt	0.52	[-2.93, 3.35]	0.34	.736	
Condition:Trial	-0.17	[-0.43, 0.07]	-1.38	.168	
NFC:Trial	0	[-0.11, 0.11]	0.07	.942	
CRT:Trial	0.02	[-0.11, 0.14]	0.37	.713	
ReflectionPrompt:Trial	-0.16	[-0.46, 0.13]	-1.08	.282	
ControlPrompt:Trial	-0.14	[-0.6, 0.33]	-0.59	.555	
Condition:NFC:Trial	0.09	[-0.01, 0.21]	1.77	.077	
Condition:CRT:Trial	0.07	[-0.04, 0.19]	1.24	.216	

Table 17: Generalized linear mixed-model for changes towards a planning strategy with higher expected game score (first 12 trials)

Fixed Effect	Estimate	95% CI	z	p	Significance
Intercept	3.19	[-1.41, 7.8]	1.36	.174	
Condition	-3.49	[-8.86, 1.89]	-1.27	.203	
Trial	-0.2	[-0.87, 0.46]	-0.6	.548	
CRT	1.29	[-0.79, 3.36]	1.22	.223	
NFC	-1	[-3.77, 1.78]	-0.7	.482	
ReflectionPrompt	5.47	[-0.55, 11.5]	1.78	.075	
ControlPrompt	16.65	[-12279.49, 12313.22]	0	1	
Condition:Trial	0.29	[-0.52, 1.1]	0.69	.488	
NFC:Trial	0.14	[-0.23, 0.51]	0.75	.453	
CRT:Trial	-0.1	[-0.53, 0.33]	-0.46	.643	
ReflectionPrompt:Trial	-0.65	[-1.45, 0.14]	-1.61	.108	
ControlPrompt:Trial	3.15	[-1945.37, 1945.16]	0	1	
Condition:NFC:Trial	0	[-0.22, 0.23]	0.02	.981	
Condition:CRT:Trial	-0.09	[-0.43, 0.25]	-0.52	.603	

Table 18: Linear mixed model for the difference in expected game score with changes in planning strategies. (first 12 trials)

Fixed Effect	Estimate	95% CI	df	t	p	Significance
Intercept	14.12	[-14.89, 43.14]	39	0.83	.41	
Condition	-17.54	[-53.54, 18.45]	39	-0.83	.409	
Trial	0.16	[-4.04, 4.36]	39	0.06	.949	
CRT	5.75	[-7.74, 19.24]	39	0.73	.47	
NFC	-12.16	[-29.53, 5.2]	39	-1.2	.238	
ReflectionPrompt	47.13	[18.42, 75.84]	39	2.81	.008	*
ControlPrompt	64.92	[0.8, 129.03]	39	1.73	.091	
Condition:Trial	1.06	[-4.41, 6.52]	39	0.33	.742	
NFC:Trial	1.6	[-0.91, 4.1]	39	1.09	.283	
CRT:Trial	-0.32	[-2.89, 2.25]	39	-0.21	.833	
ReflectionPrompt:Trial	-5.76	[-10.28, -1.23]	39	-2.18	.036	*
ControlPrompt:Trial	-9.19	[-18.88, 0.51]	39	-1.62	.113	
Condition:NFC:Trial	-0.5	[-2.17, 1.17]	39	-0.51	.61	
Condition:CRT:Trial	-0.42	[-2.4, 1.55]	39	-0.37	.717	