Adaptive planning depth in human problem solving

Mattia Eluchans[1,2], Gian Luca Lancia[1,2], Antonella Maselli[1], Marco D'Alessando[1], Jeremy Gordon[3], Giovanni Pezzulo[1]

[1] Institute of Cognitive Sciences and Technologies, National Research Council, Rome, Italy
[2] University of Rome "Sapienza", Rome, Italy
[3] University of California, Berkeley, Berkeley, CA 94704, US

Corresponding author

Giovanni Pezzulo

Institute of Cognitive Sciences and Technologies,

National Research Council,

00185 Rome, Italy

E-mail: giovanni.pezzulo@istc.cnr.it

Abstract

We humans are capable of solving challenging planning problems, but the range of adaptive strategies that we use to address them are not yet fully characterized. Here, we designed a series of problem-solving tasks that require planning at different depths. After systematically comparing the performance of participants and AI planners, we found that when facing manageable problems that require planning to a certain number of subgoals (from 1 to 6), participants make an adaptive use of their cognitive resources – namely, they tend to select an initial plan having the minimum required depth, rather than selecting the same depth for all problems. When facing more challenging problems that require planning up to 7 or 8 subgoals, participants tend to select a simpler (greedy) strategy wherein plans consider only the next 1 or 2 subgoals. Furthermore, we found a strong similarity between how different participants solve the same problems. These results support the view of problem solving as a bounded rational process, which adapts costly cognitive resources to task demands.

Keywords: planning; problem solving; cognitive search; bounded rationality

Significance statement

Problem-solving usually requires planning ahead for several steps. While previous studies assessed that we can plan multiple steps ahead during problem solving, it is unclear whether we use planning resources parsimoniously, given their significant cognitive costs. Theories of bounded rationality suggest that rather than always operating at our maximum capacity, we should adapt planning depth to task demands. Alternatively, each of us could have a preferred "planning depth" and use the same depth, regardless of the problem. By comparing the behaviour of human and AI planners during problem solving, we found that people adapt planning depth to problem complexity; but resort to simpler strategies when the problem is very hard. Our study reveals a flexible use of bounded cognitive resources during problem solving.

Introduction

Since the early days of cognitive science, researchers have asked how we solve challenging problems that engage planning abilities, such as the Tower of Hanoi and Traveling Salesman as well as popular games such as chess or go (Newell & Simon, 1972; Shallice, 1982; Wiener et al., 2009). Most cognitive theories assume that planning requires a form of cognitive *tree search* over an internal model or mental map of the task (Craik, 1943; Daw & Dayan, 2014; Hunt et al., 2021; Kolling et al., 2012; Tolman, 1948). Following this view, several planning studies in humans and other animals used tree-like tasks, but mostly focused on simple (e.g., two-step) problems which could be searched exhaustively (Akam et al., 2015; Daw et al., 2011; Hasz & Redish, 2018; Miller et al., 2017). It is still unclear how we solve more complex problems, which – given our limited resources – defy exhaustive search.

From a normative perspective, planning with limited resources could be described as a *bounded* rational processes, which balances the accuracy of the solution and the cognitive resources invested, e.g., memory and time (Bhui et al., 2021; Lieder & Griffiths, 2020; Simon, 1957). One way to lower cognitive resources is using *heuristics* to alleviate the burden of exhaustive search (Geffner & Bonet, 2013; Russell & Norvig, 1995). For example, it has been proposed that people use a pruning heuristic during mental search: if they encounter a tree node that seems unpromising, they discard the whole branch of the tree (Huys et al., 2012, 2015; Van Opheusden et al., 2017). Other heuristics consist of sampling only a few promising routes, or many routes but only up to a certain depth (Keramati et al., 2016; Pezzulo et al., 2013). Furthermore, it is possible to alleviate the burden of planning by using a hierarchical approach to split the problem into more manageable subproblems (Balaguer et al., 2016; Donnarumma et al., 2016; Ribas-Fernandes et al., 2011; Solway & Botvinick, 2012; Tomov et al., 2020) or by interleaving planning and execution; for example, plan until a certain subgoal, then revise and complete the plan along the way, as one moves toward the chosen subgoal (Geffner & Bonet, 2013; Russell & Norvig, 1995). Despite this progress, we still have incomplete knowledge of the (approximate) planning methods that humans and other animals might adopt during problem solving, as well as their neuronal underpinning (Mattar & Lengyel, 2022; Miller & Venditto, 2021; Pezzulo et al., 2019).

Another stream of research explored the limitations (e.g., the maximum depth) of our planning abilities. Various studies have shown that with sufficient time, people are able to find near-optimal solutions to challenging problems, such as the Traveling Salesman, which requires finding the shortest possible closed path that connects a fixed number of "cities" (MacGregor & Ormerod, 1996; Vickers et al., 2001). Other studies have tried to quantify planning depth in chess (Chase & Simon, 1973; De Groot, 1946) and other games, see (van Opheusden & Ma, 2019) for a recent review. Classical studies reported that top chess players can plan ahead (on average) a relatively small number of moves, between 3.6 and 5.4 and their maximum planning depth is between 6.8 and 9.1 moves (De Groot, 1946; Gobet, 1998; Saariluoma, 1995); but see (Campitelli & Gobet, 2004) for evidence that grand masters can plan ahead (on average) 13.8 moves.

It is clear then that planning – especially when done at greater depths – requires engaging a significant amount of cognitive resources. However, the idea that planning is a *bounded* rational process that balances accuracy and cognitive complexity suggests the untested hypothesis that people can flexibly adapt their planning depth to the demands of the problem. In other words, rather than using a fixed planning depth for all problems, people might use the minimum planning depth necessary to find a solution – hence showing an adaptive use of cognitive resources (Anderson, 1990).

To test this hypothesis, we asked participants to solve a series of planning problems that required finding a path to connect all the "gems" in a grid, without passing through the same node twice (though "backtracking" to un-select nodes was allowed). Participants solved the problems by "navigating" with a finger in a grid that was fully visible on their mobile phones. They had 60 seconds to solve each problem and earned more points if they solved it faster. Figure 1 shows an example problem, which requires finding a path from the home location (yellow node) through all the gems (red nodes). The six panels show six representative timesteps of the solution, with the azure line indicating the path taken (visible to the participant) and the small red dots showing the actual finger positions at different times (not visible to participants). In the example illustrated in Figure 1, the participants see on their mobile phone the configuration shown in Panel A. They first select an incorrect path towards the two gems to the right (Panel B), then they backtrack to the home location (Panels C-D) and finally select a correct path that connects all the gems (Panels E-F) – therefore solving the problem.

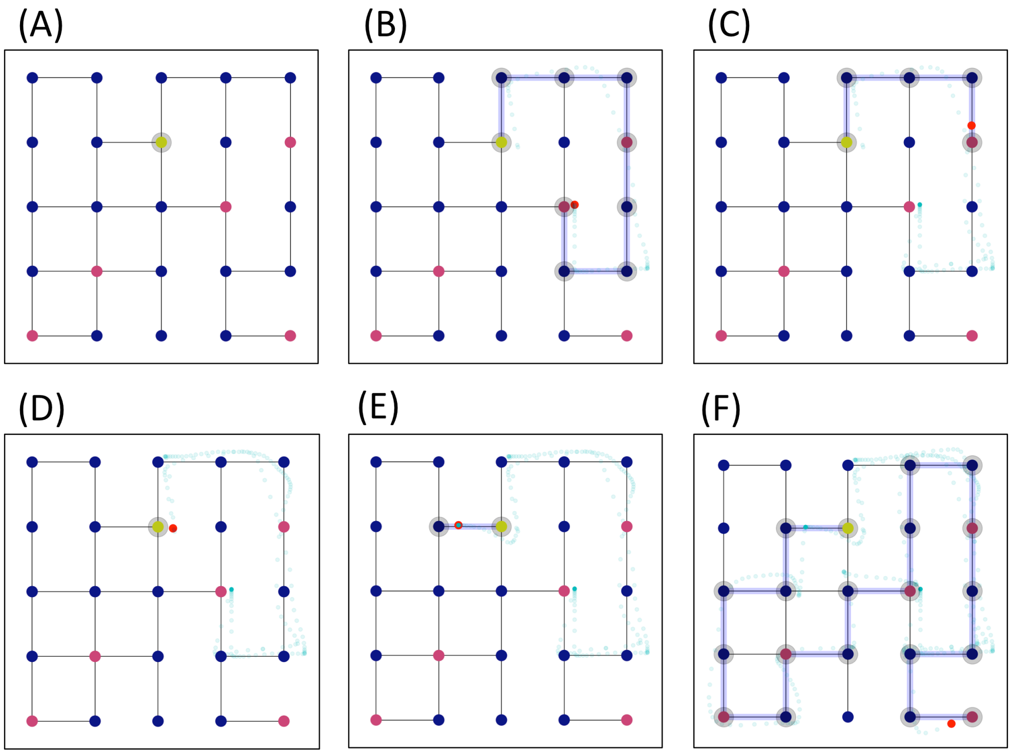

Figure 1. An example problem in the experiment. *The problem requires finding a path in the grid that starts from the home location (yellow node) and collects all the "gems" (red nodes), without passing through the same node twice. Participants solved the problem by navigating with their finger on the (fully visible) grid, on their mobile phones. The figure shows six time steps of the solution, with the azure path indicating the path taken by one of the participants, the azure dots the actual finger trajectory (sampled at 60Hz) and the small red dot the current finger position. Note that solving this particular problem requires planning 5 gems in advance. See the main text for explanation.*

Crucially, we designed problems that require different planning depths to be solved, from 1 to 8. Some problems could be solved using a "greedy" strategy to always move to the closest gem (i.e., planning depth 1), whereas other problems required optimizing the plan through the next 2 to 8 gems. For example, the problem shown in Figure 1 requires a planning depth of 5. In other words, finding the correct path (Figure 1F) requires planning 5 gems in advance and any optimal (shortest) path to a lower number of gems (from 1 to 4) makes the solution impossible.

This design allows us to study the planning depth that participants use to solve each problem, by systematically comparing their behavior with the behavior of 8 AI planners, which use planning depths from 1 to 8. We focused only on the paths chosen by participants before their first backtracks, as these could be used as a proxy for the subjects' initial planning depth.

If people always use the same planning depth (e.g., depth 3) to solve all the problems, then their behavior should always match the same AI planner (e.g., the one at depth 3) across all problems. If instead people recruit resources according to task demands, they will adapt their planning depth to the minimum depth required to solve the problem, and then their behavior should match a different planner for each set of problems; namely, the planner that uses the (set-specific) minimum depth.

To preview our results, we found that people tend to use the minimum depth required to find the solution, when the depth is between 1 and 6. In contrast, people address problems that require greater planning depth (7 or 8) by using a much lower depth – usually 1. Furthermore, we found that the above results are not an artifact of averaging across multiple participants; rather, for each particular problem, most participants tend to select the same planning depth. These results indicate that people flexibly adapt their planning depth – or how much they look ahead – to task demands, unless the problem is too hard, in which case planning costs are minimized.

Methods

Data Collection

The experiment was conducted with the support of ThinkAhead, an Android application developed to study navigational planning and problem solving[1]. We recruited 160 participants online and all gave informed consent to our procedures which were approved by the Ethical Committee of the National Research Council. In the analysis, we consider the 65 players (42 male, age = 34 +- 11 years; 19 female, age = 35 +- 10; 4 participants who preferred not to specify their gender, age = 34 +- 10 years) who tried at least 80 of the 90 problems of the experiment.

Experiment design

The experiment comprised 90 problems, each requiring participants to collect all the "gems" (i.e., colored dots) in a grid, without passing through the same node twice, in 60 seconds; see Figure 1 for an example problem. Participants were instructed that they would earn points

[1] https://play.google.com/store/apps/details?id=com.GraphGame.Conan

proportional to the time left to solve the problems and that the points were doubled in problems where the gems were red (which happened in half of the trials) compared to those where the gems were blue (in the other half of the trials). As soon as the problem was shown to the participants, the time countdown started. If the participants did not solve a problem within the deadline, it counted as a failure; participants were allowed to either complete it (without getting any points) or to skip it and pass to the next one.

We generated a range of problems that required different planning depths, from 1 to 8 (see the next section for the exact definition of planning depth). The problem grids were all unique and were generated with an average level of connectivity of .75 +- .2, which we found in a pilot study to afford a good range of planning solutions.

Before the experiment, participants performed a short practice session, in which they had to solve 4 problems, whose results were not analysed. The 90 problems were divided into 3 blocks (henceforth, "levels"), with 30 problems for each level. We varied planning demands both between and within levels; see Supplementary Figure S1 and Supplementary Table S1 for details. To vary planning demands between levels, the 3 levels were characterized by increasingly large maps and more gems to be collected, making the higher-level problems (on average) more challenging. To vary planning demands within levels, we divided each level into 3 sublevels, of 10 problems each. The 3 sublevels comprised problems that could be solved using planning depth 1, 2-3, or 4-8, respectively.

AI planners and assessment of the minimum planning depth required for each problem

We designed 8 AI planners that use a depth-first search strategy over all the simple paths that start from the current position and reach n gems, with n being the planning depth (from 1 to 8) and the only parameter of the class of algorithm (Geffner & Bonet, 2013; Russell & Norvig, 1995). For example, a "greedy" planner with $n = 1$ will look for all the simple paths (i.e., paths that do not pass twice on the same node) that reach one and only one gem. A planner with $n = 2$ will look for all the simple paths crossing only 2 gems, and so on. After the computation, the shortest path is selected. If there are many shortest paths the choice will be uniformly random among them. From the new position (i.e., the end of the last selected path), and with the nodes from the chosen paths being removed to comply with task rules, the computation of the next (partial) paths is repeated, until there are no more gems to collect, or the planner has reached a dead-end. The planners cannot backtrack, so once a dead-end is reached, the simulated trial ends. Note that since there might be various paths to choose among, multiple runs of the same agent on the same problem might have different outputs. See the Supplementary Materials for the pseudocode of the AI planning algorithms.

We used the AI planners to assess the minimum planning depth required to solve each problem. For this, we classified each problem according to the minimum value of n for which a solution could be found, from 1 to 8; for example, problems of depth 5 can only be solved by AI planners at depth 5 to 8, but not by AI planners at depth 1 to 4. This permits us to group the problems into 8 sets, with the index denoting the (minimum) planning depth required to solve them.

Similarity between participants and AI planners, for each of the 8 problem sets

The comparison of the behavior of human participants and AI planners was performed by calculating the similarity between the depth of their plans, in each of the 8 problem sets requiring minimum planning depth from 1 to 8. For this, we compared the distribution of the number of gems collected by the whole group of participants for any given problem before the first backtrack, and the distributions of the number of gems collected by 200 instances of each of the 8 AI planners (note that multiple instances of each planner are necessary to cover the full spectrum of possible solutions at a given depth with equal distance). Then, we computed the Kullback–Leibler (KL) divergences of the human distributions with respect to the distributions of each of the 8 AI planners. We selected the minimum among the KL divergences to determine the most similar AI planner, for each of the 90 problems.

Stereotypy of planning depth across participants on each problem

Furthermore, we devised a measure of stereotypy of the planning depth that the participants select to solve the same problems. A high stereotypy indicates a strong similarity between how different participants solve the same problems. Conversely, a low stereotypy indicates that participants solve the same problems in different ways. To calculate stereotypy, we first needed to assess the similarity between participants and AI planners, at the level of single participants, rather than at the group level as above. For this, we compared the number of gems collected by each participant in a given trial, before the first backtrack, with the *average* number of gems collected from 200 instances of each AI planner. We picked the AI planner whose average number of gems was the closest to that of the participant as the most similar planner. Note that while this measure of similarity is less precise than the group-level measure used above (as it does not use the full distribution of the solutions generated by participants and AI planners), it has the advantage that it can be used at the individual level. Then, to estimate the variability of planning depth on each problem, we computed the mode of the most similar AI planner and then estimated the proportion of participants who chose their planning depth according to the mode. Finally, we averaged this number across the 8 sets of problems sharing the same minimum planning depth.

Results

Below we report the results of the comparison of human participants and AI planners, using the two measures of *similarity* and *variability* described in the Methods, to characterize behaviour before the first backtrack. Note that in the analyses reported in the main text, we aggregate the results across the three levels and for red and blue gems, because we did not find significant differences between them (see Supplementary Figures S2 and S3 for the results of separate analyses). See also the Supplementary Materials for additional analyses of participants' success probability (i.e., the probability that they solved the problems before the deadline of 60 seconds), problem completion time (i.e., the average time they needed to complete the problem, in seconds) and total number of backtracks that executed during the experiment.

Participants' initial planning depth is adaptive and matches task demands, unless the problem is too hard.

We first assessed the similarity between the initial plans of participants and AI planners, for each of the 8 problem sets. The results of this analysis are shown in Figure 2. The 8 rows group the problems according to the minimum planning depth required to solve them, from 1 to 8 (e.g., the first row indicates the problems that could be solved by a planner at depth 1). For each row, the figure shows the probability distribution over the AI planners that best match participants' behavior, in the row-specific set of problems. The element identified by the *i-th* row and *j-th* column of the matrix represents the probability (computed as a frequency) that for a problem that required a minimum planning depth equal to *i* the most similar depth to humans was equal to *j*. For example, consider that in the fifth row, most of the probability mass is concentrated on the fifth column. This indicates that in most problems requiring planning depth 5, the behavior of AI planners with depth 5 provides the best match with the participants' behavior.

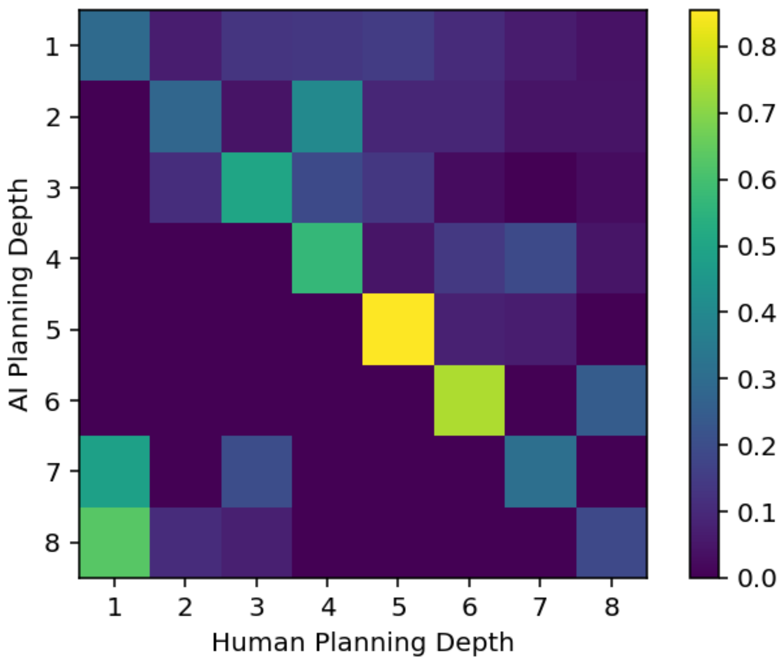

Figure 2. *Comparison of the initial planning depth of participants and AI planners. The row index indicates the minimum planning depth (from 1 to 8) required to solve each set of problems. The columns indicate the planning depth selected by the participants. For each row, the figure shows the probability distribution over the AI planners that are most similar to the human participants, in the row-specific set of problems – with the cells having the highest probabilities indicating the best matching. For problems 1 to 6, much of the probability mass is on the diagonal or slightly above it, indicating that participants tended to select the minimum planning depth required to solve these problems, or a slightly greater depth. Rather, for problems 7 and 8, much of the probability mass is on the first column, indicating that participants tended to use a greedy strategy to plan at depth 1 to solve these problems.*

Our results indicate that in the 6 simpler sets of problems (rows 1 to 6), most of the probability mass lies on the diagonal, or slightly above it. This implies that in these problems, the participants tend to select an initial plan at the minimum required depth, or a slightly greater depth. In the two more challenging sets of problems (rows 7 and 8), however, participants are best aligned with greedy planners that use depth 1. We obtained similar results when we analysed the problems separately for the 3 levels of the experiment (Supplementary Figure S2), separately for red or blue gems (Supplementary Figure S3), and when we grouped them into classes reflecting their minimum required depth and maximum number of gems (Supplementary Figures S4 and S5). These control analyses indicate that our main findings do not depend on map size (which varies across levels and class groups) or incentives (which are different for red and blue gems).

Strong similarity between how different participants solve the same problems

We next assessed the stereotypy of participants' initial planning depth, in the 8 problem sets. The results of the analysis are shown in Figure 3. There is a decreasing linear trend of the behavioral stereotypy as the planning depth increases (orange line), which was estimated through Pearson test (coefficient: -0.94, p-value = 3*10^-4). The figure shows that the proportion of problems that satisfy the behavior stereotypy criterion (blue dot-dash line) is very high and beyond chance level (black dot-dash line), across all problem sets. Indeed, the slope of participants' stereotypy (m = -3.5*10^-2, std = 4.9*10^-3) was significantly smaller than the slope of the random case (m = -6,4*10^-3, Z-test = -4,89, p-value <10^-5). We confirmed this statement by evaluating the statistical significance of the interaction between the depth and the group (Participants vs. Random case) with an ANOVA test (F = 2.98, p-value = 5.8*10^-3).

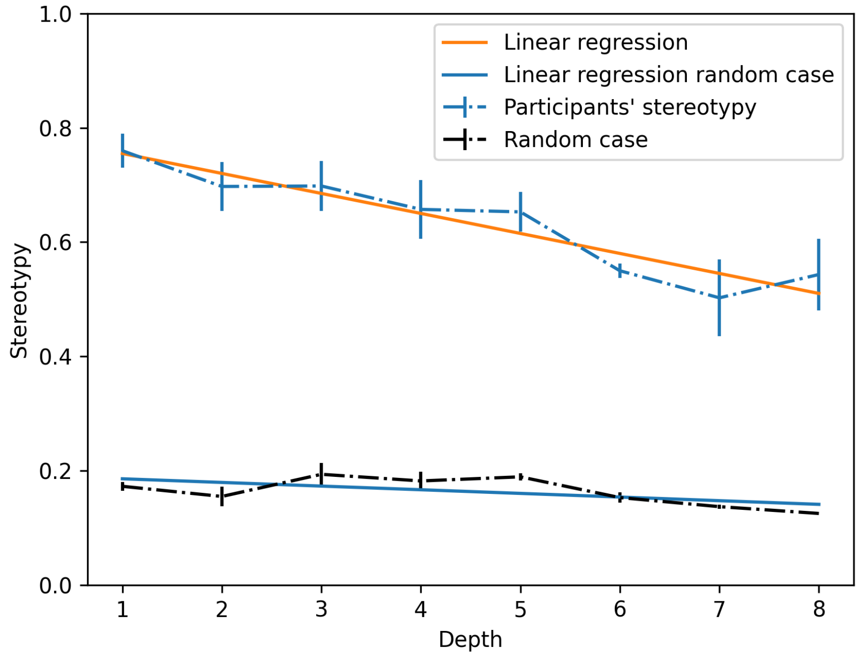

Figure 3. *Stereotypy of participants' initial planning depth, in the 8 problem sets. The figure shows the mean stereotypy across problems having minimum planning depth from 1 to 8. The blue dotted bar shows the data trend, with error bars indicating the standard error of the mean, obtained by averaging problem stereotypy in each problem set. The orange line indicates the linear fit. The black dashed line indicates chance level. Note that since the number of possible planners is bounded by the maximum number or gems in the problem, the expectation for the null case (random across the possible depths) varies across the 8 problem sets. See the main text for explanation.*

Figure 4 further unpacks the results illustrated in Figure 3, by showing the results for each participant and problem. Each vertical block in Figure 4 indexes one of the 8 sets of problems having planning depth (D) 1 to 8, each row indexes one participant and the colors indicate the depth of the planner that best resembles the performance of each participant in each problem (white dots are missing points, e.g., skipped problems). The figure helps appreciate that many participants tend to select the same planning depth to address the same problem, as evident by the numerous vertical bands of Figure 3. Notably, this happens also in "misleading" trials, such as problems 10 and 23 in the first column of Figure 3, which were addressed using high planning depths, despite its minimum depth was 1.

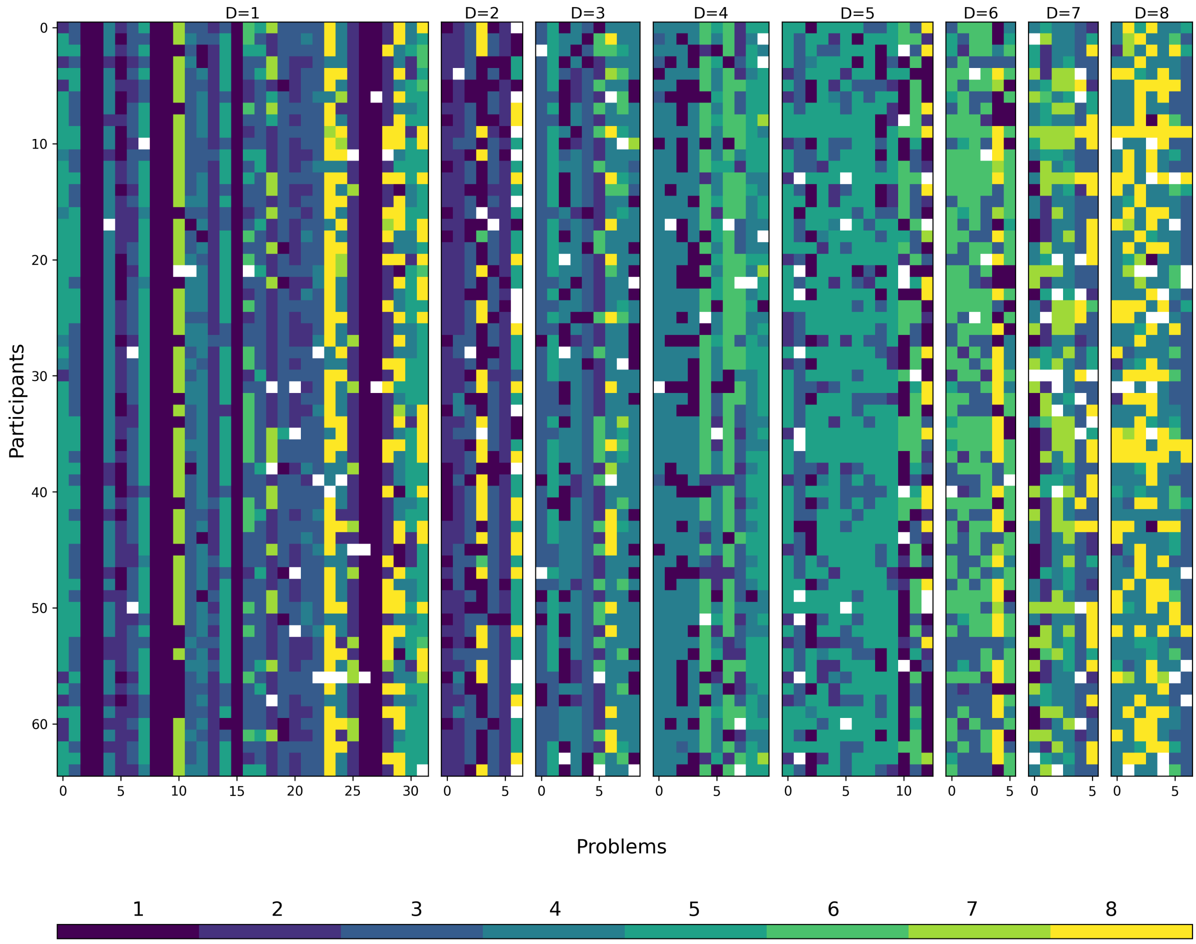

Figure 4. *Variability of the planning depth. Each row corresponds to a different participant (y-axis) and each column to a specific problem (x-axis), so each square is a different trial. Problems are grouped according to the minimum required planning depth (D, in the figure), ranging from 1 to 8. The color bar, going from 1 (dark blue) to 8 (yellow), assigns to each trial the color of the most similar AI planner. White squares are missing or discarded trials. The resulting vertical "bands" of the same colors suggest the presence of problems for which several participants used the same planning depth (see the main text for explanation).*

Discussion

Since the early days of cognitive science, researchers have asked how we solve complex planning problems that defy the exhaustive evaluation of all possible choices (Newell & Simon, 1972). It is commonly assumed that planning is a form of cognitive tree search using a mental map (Craik, 1943; Daw & Dayan, 2014; Hunt et al., 2021; Kolling et al., 2012; Tolman, 1948). However, except in the simplest cases, exhaustive search is infeasible, hence pointing to *bounded* forms of planning that adapt cognitive resources to task demands (Bhui et al., 2021; Lieder & Griffiths, 2020; Simon, 1957).

In this study, we asked whether participants adapt their planning resources to the demands posed by problems of different complexity. For this, we asked participants to solve a series of planning problems that required finding a path to connect all the "gems" in a grid, without passing through the same node twice. Crucially, solving the problems required different levels of planning depth, which was unknown to the participants but which we could estimate with the help of AI planners. By systematically comparing the behavior of participants and AI planners, we aimed to assess whether participants kept their (initial) planning depth fixed across problems of changed it adaptively, as a function of problem complexity.

The main contribution of this study is a demonstration that when facing relatively manageable problems, participants' behaviour is compatible with a flexible adaptation of their initial planning depth to the minimum depth required to solve the problem. In contrast, participants tend to use a simpler, "greedy" strategy to deal with more challenging problems, which would have required an excessive amount of cognitive resources. These results therefore support an adaptive use of cognitive resources during problem solving (Anderson, 1990). Furthermore, we found a strong correlation between how different participants solve the same problem, which permits ruling out the possibility that our results are an artefact of averaging across multiple participants.

This study adds to a large literature showing that when facing challenging problems that eschew exhaustive search, people adopt various simplifications. For example, people prune unpromising branches of the search tree (Huys et al., 2012, 2015) and reduce tree search under time pressure (Van Opheusden et al., 2017). An emerging idea is that these (and other) simplifications might be "rational", in the sense that they entail a flexible and efficient use of limited resources– i.e., *bounded* or *resource-rational* planning (Callaway et al., 2018, 2022). For example, during problem solving, people might spend more time planning ahead when the benefits of investing cognitive resources are greater (Russek et al., 2022).

In keeping with this perspective, our study shows that when participants address manageable problems, they use their cognitive resources efficiently, by selecting initial plans that have the minimum required depth, or a slightly greater depth (Figure 1). These initial plans therefore tend to use the appropriate depth, in keeping with previous reports that the initial moves of participants during problem solving are usually "good enough" (Klein et al., 1995). Note also that there is a slight asymmetry in the "matrix" shown in Figure 1: the divergence from minimum planning depth (the diagonal) is usually realized through a higher-then-needed investment of resources. However, the above pattern of results breaks down when addressing the two more challenging sets of problems – because to address these problems, participants often use a much lower planning depth. Note that neither the levels nor the economic incentives (i.e., the different values of red and blue gems) significantly modulate planning depth, plausibly because the task is challenging and participants might always perform at ceiling level.

Furthermore, we found a strong similarity between how different participants solve the same problem. This stereotypy was partially unexpected, given that previous studies of navigational planning showed a greater variability of strategies (Krichmar & He, 2021; Lancia et al., 2023). However, please note that our experiment required lower memory demands than the previous studies, since the participants could observe the whole problem graph. It is possible that part of the variability reported in the previous studies was due to incorrect map memorization, not planning; but this hypothesis remains to be tested in future studies.

Taken together, these results indicate that during problem solving, people make an adaptive use of their cognitive resources by selecting the appropriate level of planning depth – or a much simpler (and cheaper) planning strategy if the problem is too challenging.

This study has various limitations that need to be addressed in future work. First, this study indicates that participants adapt their initial planning depth to task demands, but does not clarify how they do that. There are multiple alternative strategies that could explain our findings. For example, before navigation begins, participants might use a "gist" of the maze to decide planning depth – or how much cognitive resources they would need to invest to plan ahead – and then use such planning depth to find a solution (note that selecting an appropriate planning depth does not necessarily entail solving a particular problem, because there are usually several alternative paths at the same planning depth). In the planning literature, there is a key distinction between an "encoding" phase in which participants form a mental representation of the problem and the "planning" phase, in which they form a plan based on the mental representation. Previous results suggest that during the encoding phase, participants might form simplified mental representations of the problems, which omit many details (Ho et al., 2022). It is possible that this simplified (or gist) representation could be sufficient to guide the selection of an appropriate planning depth, but future studies are required to understand whether and how this is possible. Alternatively, before navigation begins, participants might start searching for a solution at low planning depth and then increase the depth progressively, until they find a (satisficing) solution. This procedure, which is akin to "iterative deepening depth-first search" in AI (Korf, 1985) and to algorithms that start from an initial "cheap" plan and progressively refine it (Lancia et al., 2023; Todorov, 2009), would impose a filter on insufficient planning depth, which could result in the over-investing bias seen across problems 1 to 6. In other words, it could be faster to recognize that your strategy is too poor for the problem rather than recognizing that you are investing too much. However, when the required planning depth is deemed excessive, participants might give up and rather select the least costly initial plan (i.e., a plan at low depth) – and possibly revise it along the way. The plausibility of these or alternative strategies for adapt planning depth remains to be fully investigated in future studies.

A related limitation is that this study does not disambiguate the "algorithm" used by the brain to solve the problems. When establishing a similarity between the AI planners and human participants, we are not necessarily claiming that human participants use depth-first planning, but only that they appear to adapt their planning depth across problems – regardless of mechanism (e.g., a different planner for each problem, "iterative deepening depth-first search", or other methods). Having said this, our specific setting, in which the problem maps are novel and fully visible, epitomizes the use of model-based planners (Daw & Dayan, 2014; Dickinson & Balleine, 1994; Dolan & Dayan, 2013; Friston et al., 2016; Parr et al., 2022). While in principle the problems considered in this experiment could be solved using model-free and successor representation algorithms that dispense from planning (Dayan, 1993; Gershman, 2018; Sutton

& Barto, 1998), these are unlikely candidates, since they would require an extensive learning phase, whereas in our experiment the participants never see the same map twice. Future studies might compare more directly various planning (or even non-planning) methods. Apart for those discussed above, another relevant class of algorithms is Monte Carlo planning. These algorithms offer an approximate solution to the problem of sampling in large or continuous state spaces (Silver & Veness, 2010) and have been linked to human cognitive search (Jensen et al., 2023; Mattar & Lengyel, 2022; Pezzulo et al., 2019), but would behave similarly to depth-first planners in our (relatively small scale) problems. Another possibility is using hierarchical planners, which split large problems into smaller, more manageable ones (Donnarumma et al., 2016; Solway et al., 2014; Tomov et al., 2020). Given that the focus of this paper is on the first plans that people form, not on their entire plans, the use of hierarchical planners seems less compelling; but future studies could use hierarchical planners to extend the results of this study to the entire plan selected by participants.

This leads us to another limitation of this study: the fact that it only focuses on the initial planning phase. Focusing on the first part of the plan is meaningful, since previous research has established that the initial moves of participants during problem solving are revelatory of their strategy (Klein et al., 1995). Furthermore, from a methodological perspective, considering only the first part of the plan drastically simplifies the assessment of planning depth, since it does not require making particular assumptions about the algorithm used to decide (for example) when to backtrack or whether to change plan or plan depth along the way, as some AI planners do (Weld, 1994). We hope to address these ongoing changes-in-planning dynamics in future studies.

Finally, another limitation of the study is that by restricting our analysis to participants who downloaded the game and completed at least 80 problems, we could have selected those having sufficient skill and/or engagement levels. Previous studies with a much larger pool of participants reported significant individual differences in navigation ability (Coutrot et al., 2022), suggesting that weaker navigators might not show the same adaptive use of resources that we report here. Therefore, the possible differences in adaptive planning depth between good and weak navigators remains to be tested in future studies.

Acknowledgements

This research received funding from the European Union's Horizon 2020 Framework Programme for Research and Innovation under the Specific Grant Agreement No. 945539 (Human Brain Project SGA3) to GP and the European Research Council under the Grant Agreement No. 820213 (ThinkAhead). The GEFORCE Quadro RTX6000 and Titan GPU cards used for this research were donated by the NVIDIA Corporation. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Software note

Giovanni Pezzulo devised the ThinkAhead App concept. The App has been implemented in Unity by Antonella Maselli on a general-purpose architecture for maps navigation games previously implemented by Massimiliano Schembri. Marco D'Alessandro and Jeremy Gordon contributed with setting the online communication with a dedicated database used to store users' progresses and solutions. Gian Luca Lancia edited the video tutorial. Mattia Eluchans prepared the list of problems shown to participants for the current study.

References

Akam, T., Costa, R., & Dayan, P. (2015). Simple Plans or Sophisticated Habits? State, Transition and Learning Interactions in the Two-Step Task. *PLOS Computational Biology*, *11*(12), e1004648. https://doi.org/10.1371/journal.pcbi.1004648

Anderson, J. R. (1990). *The adaptive character of thought* (pp. xii, 276). Lawrence Erlbaum Associates, Inc.

Balaguer, J., Spiers, H., Hassabis, D., & Summerfield, C. (2016). Neural mechanisms of hierarchical planning in a virtual subway network. *Neuron*, *90*(4), 893–903.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*, 1–48. https://doi.org/10.18637/jss.v067.i01

Bhui, R., Lai, L., & Gershman, S. J. (2021). Resource-rational decision making. *Current Opinion in Behavioral Sciences*, *41*, 15–21. https://doi.org/10.1016/j.cobeha.2021.02.015

Callaway, F., Lieder, F., Das, P., Gul, S., Krueger, P. M., & Griffiths, T. (2018). A resource-rational analysis of human planning. *CogSci*.

Callaway, F., van Opheusden, B., Gul, S., Das, P., Krueger, P. M., Griffiths, T. L., & Lieder, F. (2022). Rational use of cognitive resources in human planning. *Nature Human Behaviour*, 1–14.

Campitelli, G., & Gobet, F. (2004). Adaptive expert decision making: Skilled chess players search more and deeper. *ICGA Journal*, *27*(4), 209–216.

Chase, W. G., & Simon, H. A. (1973). Perception in chess. *Cognitive Psychology*, *4*, 55–81.

Coutrot, A., Manley, E., Goodroe, S., Gahnstrom, C., Filomena, G., Yesiltepe, D., Dalton, R., Wiener, J. M., Hölscher, C., Hornberger, M., & others. (2022). Entropy of city street networks linked to future spatial navigation ability. *Nature*, *604*(7904), 104–110.

Craik, K. (1943). *The Nature of Explanation*. Cambridge University Press.

Daw, N. D., & Dayan, P. (2014). The algorithmic anatomy of model-based evaluation. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *369*(1655), 20130478.

Daw, N. D., Gershman, S. J., Seymour, B., Dayan, P., & Dolan, R. J. (2011). Model-based influences on humans' choices and striatal prediction errors. *Neuron*, *69*(6), 1204–1215. https://doi.org/10.1016/j.neuron.2011.02.027

Dayan, P. (1993). Improving generalization for temporal difference learning: The successor representation. *Neural Computation*, *5*(4), 613–624.

De Groot, A. D. (1946). Thought and choice in chess. In *Thought and Choice in Chess*. De Gruyter Mouton.

Dickinson, A., & Balleine, B. (1994). Motivational control of goal-directed action. *Animal Learning & Behavior*, *22*(1), 1–18. https://doi.org/10.3758/BF03199951

Dolan, R. J., & Dayan, P. (2013). Goals and habits in the brain. *Neuron*, *80*(2), 312–325. https://doi.org/10.1016/j.neuron.2013.09.007

Donnarumma, F., Maisto, D., & Pezzulo, G. (2016). Problem Solving as Probabilistic Inference with Subgoaling: Explaining Human Successes and Pitfalls in the Tower of Hanoi. *PLOS Computational Biology*, *12*(4), e1004864. https://doi.org/10.1371/journal.pcbi.1004864

Friston, K. J., FitzGerald, T., Rigoli, F., Schwartenbeck, P., & Pezzulo, G. (2016). Active Inference: A Process Theory. *Neural Computation*, 1–49. https://doi.org/10.1162/NECO_a_00912

Geffner, H., & Bonet, B. (2013). *A Concise Introduction to Models and Methods for Automated Planning*. Morgan & Claypool Publishers.

Gershman, S. J. (2018). The Successor Representation: Its Computational Logic and Neural Substrates. *Journal of Neuroscience*, *38*(33), 7193–7200. https://doi.org/10.1523/JNEUROSCI.0151-18.2018

Gobet, F. (1998). *Chess players' thinking revisited*.

Hasz, B. M., & Redish, A. D. (2018). Deliberation and Procedural Automation on a Two-Step Task for Rats. *Frontiers in Integrative Neuroscience*, *12*. https://www.frontiersin.org/articles/10.3389/fnint.2018.00030

Ho, M. K., Abel, D., Correa, C. G., Littman, M. L., Cohen, J. D., & Griffiths, T. L. (2022). People construct simplified mental representations to plan. *Nature*, *606*(7912), 129–136.

Hunt, L. T., Daw, N. D., Kaanders, P., MacIver, M. A., Mugan, U., Procyk, E., Redish, A. D., Russo, E., Scholl, J., Stachenfeld, K., Wilson, C. R. E., & Kolling, N. (2021). Formalizing planning and information search in naturalistic decision-making. *Nature Neuroscience*, *24*(8), 1051–1064. https://doi.org/10.1038/s41593-021-00866-w

Huys, Q. J. M., Eshel, N., O'Nions, E., Sheridan, L., Dayan, P., & Roiser, J. P. (2012). Bonsai trees in your head: How the pavlovian system sculpts goal-directed choices by pruning decision trees. *PLoS Computational Biology*, *8*(3), e1002410. https://doi.org/10.1371/journal.pcbi.1002410

Huys, Q. J. M., Lally, N., Faulkner, P., Eshel, N., Seifritz, E., Gershman, S. J., Dayan, P., & Roiser, J. P. (2015). Interplay of approximate planning strategies. *Proceedings of the National Academy of Sciences*, *112*(10), 3098–3103. https://doi.org/10.1073/pnas.1414219112

Jensen, K. T., Hennequin, G., & Mattar, M. G. (2023). A recurrent network model of planning explains hippocampal replay and human behavior. *BioRxiv*, 2023–01.

Keramati, M., Smittenaar, P., Dolan, R. J., & Dayan, P. (2016). Adaptive integration of habits into depth-limited planning defines a habitual-goal–directed spectrum. *Proceedings of the National Academy of Sciences*, *113*(45), 12868–12873. https://doi.org/10.1073/pnas.1609094113

Klein, G., Wolf, S., Militello, L., & Zsambok, C. (1995). Characteristics of skilled option generation in chess. *Organizational Behavior and Human Decision Processes*, *62*(1), 63–69.

Kolling, N., Behrens, T. E. J., Mars, R. B., & Rushworth, M. F. S. (2012). Neural mechanisms of foraging. *Science (New York, N.Y.)*, *336*(6077), 95–98. https://doi.org/10.1126/science.1216930

Korf, R. E. (1985). Depth-first iterative-deepening: An optimal admissible tree search. *Artificial Intelligence*, *27*(1), 97–109.

Krichmar, J. L., & He, C. (2021). Importance of path planning variability: A simulation study. *Topics in Cognitive Science*.

Lancia, G. L., Eluchans, M., D'Alessandro, M., Spiers, H. J., & Pezzulo, G. (2023). Humans account for cognitive costs when finding shortcuts: An information-theoretic analysis of navigation. *PLOS Computational Biology*, *19*(1), e1010829.

Lieder, F., & Griffiths, T. L. (2020). Resource-rational analysis: Understanding human cognition as the optimal use of limited computational resources. *Behavioral and Brain Sciences*, *43*. https://doi.org/10.1017/S0140525X1900061X

MacGregor, J. N., & Ormerod, T. (1996). Human performance on the traveling salesman problem. *Perception & Psychophysics*, *58*(4), 527–539.

Mattar, M. G., & Lengyel, M. (2022). Planning in the brain. *Neuron*.

Miller, K. J., Botvinick, M. M., & Brody, C. D. (2017). Dorsal hippocampus contributes to model-based planning. *Nature Neuroscience*, *20*(9), 1269–1276. https://doi.org/10.1038/nn.4613

Miller, K. J., & Venditto, S. J. C. (2021). Multi-step planning in the brain. *Current Opinion in Behavioral Sciences*, *38*, 29–39.

Newell, A., & Simon, H. A. (1972). *Human problem solving*. Prentice-Hall.

Parr, T., Pezzulo, G., & Friston, K. J. (2022). *Active Inference: The Free Energy Principle in Mind, Brain, and Behavior*. MIT Press.

Pezzulo, G., Donnarumma, F., Maisto, D., & Stoianov, I. (2019). Planning at decision time and in the background during spatial navigation. *Current Opinion in Behavioral Sciences*, *29*, 69–76.

Pezzulo, G., Rigoli, F., & Chersi, F. (2013). The Mixed Instrumental Controller: Using Value of Information to combine habitual choice and mental simulation. *Frontiers in Cognition*, *4*, 92. https://doi.org/10.3389/fpsyg.2013.00092

Ribas-Fernandes, J. J., Solway, A., Diuk, C., McGuire, J. T., Barto, A. G., Niv, Y., & Botvinick, M. M. (2011). A neural signature of hierarchical reinforcement learning. *Neuron*, *71*(2), 370–379.

Russek, E., Acosta-Kane, D., van Opheusden, B., Mattar, M. G., & Griffiths, T. (2022). *Time spent thinking in online chess reflects the value of computation*.

Russell, S. J., & Norvig, P. (1995). *Artificial intelligence a modern approach*. Prentice Hall.

Saariluoma, P. (1995). *Chess players' thinking: A cognitive psychological approach*. Psychology Press.

Shallice, T. (1982). Specific Impairments of Planning. *Royal Society of London Philosophical Transactions Series B*, *298*, 199–209.

Silver, D., & Veness, J. (2010). Monte-Carlo planning in large POMDPs. *Advances in Neural Information Processing Systems*, *23*.

Simon, H. A. (1957). *Models of man—Social and rational.* John Wiley and Sons.

Solway, A., & Botvinick, M. M. (2012). Goal-directed decision making as probabilistic inference: A computational framework and potential neural correlates. *Psychol Rev*, *119*(1), 120–154. https://doi.org/10.1037/a0026435

Sutton, R. S., & Barto, A. G. (1998). *Reinforcement Learning: An Introduction*. MIT Press.

Todorov, E. (2009). Efficient computation of optimal actions. *Proc Natl Acad Sci U S A*, *106*(28), 11478–11483. https://doi.org/10.1073/pnas.0710743106

Tolman, E. C. (1948). Cognitive maps in rats and men. *Psychological Review*, *55*, 189–208.

Tomov, M. S., Yagati, S., Kumar, A., Yang, W., & Gershman, S. J. (2020). Discovery of hierarchical representations for efficient planning. *PLoS Computational Biology*, *16*(4), e1007594.

van der Kooij, K., Hennink, T., & others. (2021). Motivation as a function of success frequency. *Motivation and Emotion*, *45*(6), 759–768.

Van Opheusden, B., Galbiati, G., Bnaya, Z., Li, Y., & Ma, W. J. (2017). A computational model for decision tree search. *CogSci*.

van Opheusden, B., & Ma, W. J. (2019). Tasks for aligning human and machine planning. *Current Opinion in Behavioral Sciences*, *29*, 127–133. https://doi.org/10.1016/j.cobeha.2019.07.002

Vickers, D., Butavicius, M., Lee, M., & Medvedev, A. (2001). Human performance on visually presented traveling salesman problems. *Psychological Research*, *65*(1), 34–45.

Weld, D. S. (1994). An Introduction to Least-Commitment Planning. *AI Magazine*, *15*(4), 27–61.

Wiener, J. M., Ehbauer, N. N., & Mallot, H. A. (2009). Planning paths to multiple targets: Memory involvement and planning heuristics in spatial problem solving. *Psychological Research PRPF*, *73*(5), 644–658.