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 planning models, we found that when facing problems that require planning to a certain number of subgoals (from 1 to 8), 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. 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

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". Different explanations of the Traveling Salesman are based on the idea that participants could solve complex problems by using local planning (Vickers et al., 2001) or by avoiding planning ahead in depth and instead valuating the global perceptual properties of the problem to generate their solution (MacGregor & Ormerod, 1996). Other studies have tried to quantify planning depth in chess (Chase & Simon, 1973; De Groot, 1946) and other games (Krusche et al., 2018), 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) and a recent large-scale study reported that planning depth in games increases with expertise up to a ceiling at a similar level (van Opheusden et al., 2023); but see (Campitelli & Gobet, 2004) for evidence that chess grand masters can plan ahead (on average) 13.8 moves.

Planning – especially when done at greater depths – requires engaging significant cognitive resources and hence a key question is whether people make an adaptive or resource-rational (Callaway et al., 2018, 2022) use of these resources. Deciding the appropriate planning depth for a particular problem can be seen as an instance of metaplanning: namely, the formation of a plan about how to approach the planning problem (Alterman, 1988; Gershman et al., 2015; Hayes-Roth & Hayes-Roth, 1979; Kuperwajs et al., 2024; Wilensky, 1981). One possibility is that people are sensitive to the planning demands of different problems and flexibly adapt their planning depth to the minimum depth necessary to solve them. An alternative possibility is that people are insensitive to planning demands and use a fixed planning depth to solve all the problems.

To disambiguate between these hypotheses, 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), solving the problem.

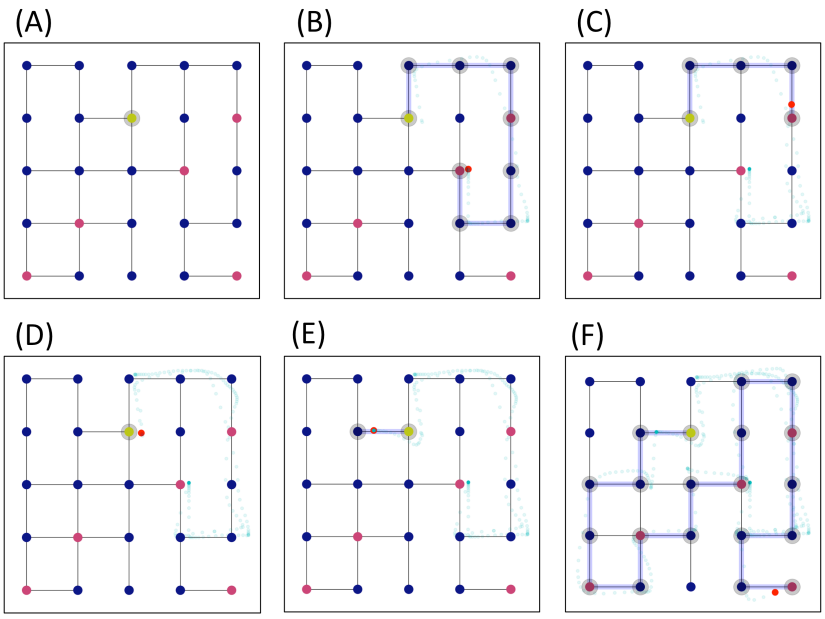

Figure 1. An example problem in the experiment. *The problem requires finding a path in the grid that starts from the home location (yellow node) and collects all the "gems" (red nodes), without passing through the same node twice. Participants solved the problem by navigating with their finger on the (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.

Crucially, we designed problems that require different planning depths to be solved. The concept of planning depth has not a unique definition, but is in general used to refer to the property of "how much further from my current state am I thinking of". In this work, we anchor the planning depth to the number of gems (from 1 to 8) that a person considers collecting in the current plan. Some problems could be solved using a "greedy" strategy to always move to the closest gem (i.e., planning depth 1). However, in other problems, following the "greedy" strategy leads to a dead end. To solve these problems, it is necessary to select a plan to move to the next 2 to 8 closest gems (i.e., planning depths 2 to 8). For example, the problem shown in Figure 1 requires a planning depth of 5. This means that finding a(ny) solution requires planning 5 gems in advance and any planner that only considers a smaller number of gems would fail. Henceforth, a problem is said to require planning depth *n* if *n* is the minimum number of closest gems that an agent has to take into account in the plan, in order to be able to solve the problem.

This design allows us to study the planning depth that participants use to solve each problem, by systematically comparing their behaviour with the behaviour of 8 variants of the same planner, which only differ for the depth *n* parameter (with depth *n* ranging from 1 to 8). In our analysis, we focused solely on the paths chosen by participants before they made their first backtracks, as these paths serve as valuable indicators of their initial plans. This approach also allows for easier alignment with findings from studies on other games like chess, where movements cannot be reversed.

If people always use the same planning depth (e.g., depth 3) to solve all the problems, then their behaviour should always match the same planning model (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 behaviour 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 to any specific problem, indicating that they flexibly adapt their planning resources to the situation.

Methods

Data Collection

The experiment was conducted with the support of ThinkAhead, an Android application developed to study navigational planning and problem solving (https://play.google.com/store/apps/details?id=com.GraphGame.Conan). 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. Participants were free to leave the experiment at any moment. 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, although they did not necessarily solve all of them.

Experiment design

The experiment comprised 90 problems, each requiring participants to collect all the "gems" (i.e., colored dots) in a grid, without passing twice through the same node, in 60 seconds; see Figure 1 for an example problem and Figure S1 for other example problems of the three levels and planning depths. Participants were instructed that they would earn points 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 edges density of 0.75 +- 0.2 (i.e., on average, 25% of all the possible edges where removed), 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 interleaved blocks (henceforth, "levels"), with 30 problems for each level. We varied planning demands both within and between levels; see Table 1 and Supplementary Figure S1. 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, 2-4, or 5-8, respectively; see Table 1. Note that the levels have (on average) increasing difficulty but include a fixed proportion of easier trials (i.e., requiring lower planning depth). This design permits increasing the challenge and simultaneously maintaining a high frequency of success across levels, which is important for intrinsic motivation (van der Kooij et al., 2021).

Table 1. Experiment design. *The experiment is structured into 3 levels of increasing (average) difficulty, each including 30 problems. See the main text for explanation.*

	Level 1	**Level 2**	**Level 3**
Sublevel 1	Depth = 1 Size = 5x5 Gems = 4-5	Depth = 1 Size = 6x6 Gems =6-7	Depth = 1 Size = 6x8 Gems = 8
Sublevel 2	Depth = 2-4 Size = 5x5 Gems = 4-5	Depth = 2-4 Size = 6x6 Gems = 6-7	Depth = 2-4 Size = 6x8 Gems = 8
Sublevel 3	Depth >4 Size = 5x5 Gems = 4-5	Depth >4 Size = 6x6 Gems = 6-7	Depth >4 Size = 6x8 Gems = 8

Assessment of the minimum planning depth required for each problem

To assess participants' planning depth during the experiment, we designed a planner that uses a depth-first search strategy over all the simple paths that start from the current position and reaches *n* gems. In our analyses, we consider 8 variants of the same planner, which only differ for the *n* parameter, which represents planning depth (from 1 to 8) (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 planner on the same problem might have different outputs. Please note that the use of the depth-first strategy for the search of the paths was just a method to enumerate all the possible paths (compatible with the constraints of the task and the instantaneous configuration) and any other exhaustive search algorithm (e.g., breadth-first) would produce the same output, hence our choice of a planning method does not necessarily entail the hypothesis that people use depth-first search. See the Supplementary Materials for the pseudocode of the planning algorithm.

We used the 8 variants of the planner 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 with non-null probability, from 1 to 8. For example, problems of depth 5 can only be solved by a planner with $n = 5$ (or possibly more), but not by a planner with $n = 1$ to 4.

This permits us to group the problems into 8 groups, with the index denoting the minimum planning depth required to solve them. Note that the fact that a problem can be solved by a planner with depth *n* does not imply that deeper planners can necessarily solve the same problem, or that planners at different depths find the same solution to the same problem. Let's suppose, for example, that a solution to a problem exists for a planning depth of 5. This solution reflects the shortest path connecting 5 gems, but there is no guarantee that a planner at depth 6 (or higher), which looks for the shortest way of connecting 6 gems, would find the same solution. This might be the case, for example, when a longer path to the first 5 gems is needed in order to reach the sixth one with a globally shorter solution. Notable, the fact that planners having different depths could solve the same problem in different ways permitted us to estimate participants' depths more carefully (see Supplementary Figures S5 - S12).

Analysis of success probability, problem completion time and total number of backtracks

We used multiple logistic and linear regressions and the R library lme4 (Bates et al., 2015) to analyse 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 they executed during

the experiment. The 3x3 design considers the 3 levels of the experiment (Levels 1-3) and the 3 planning depths (DepthID 1-3, corresponding to depth 1, depths 2-4, and depths 5-8, respectively).

Similarity between participants and planning models, for each of the 8 problem groups

The comparison between human participants and planning models involved assessing the similarity in planning depth across the eight problem groups that required minimum planning depths ranging from 1 to 8. For each of the 90 problems, we determined the likelihood of the data (number of gems collected by the participants before the first backtrack) given the models (the eight planning models). To ensure comprehensive coverage of potential solutions at varying depths, we utilized 500 instances of each planner for every problem. Subsequently, we identified which planner, with a planning depth ranging from 1 to 8, exhibited the highest probability (i.e., was the 'winner') within each problem group.

Results

Participants are sensitive to both problem level and planning depth, but in different ways

We assessed the statistical significance of the effects of the experimental design (Level x DepthID) on the success probability (i.e., the probability that they solved the problems before the deadline of 60 seconds). For this, we used a general linear mixed-effect model with Level and DepthID and their interaction as fixed-effects while accounting for subject identity variability as a random effect. The values of the coefficients of the model are reported in Table 2, with their errors and significance. We were able to assess the statistical significance of the different terms of the full model by comparing it against reduced models that excluded some fixed effect. This shows significance for the main effect of the DepthID and the Level terms, while no interaction was found. Also, the significance of the use of the random term was assessed comparing the full model with one without the random term: ($\chi^2(1, 5697) = 112$, p-value $< 2.1*10^{-16}$).

Table 2. Success probability regression. *A general linear mixed-effect model is used to fit success probability. Significance code (0< *** <0.001; 0.001** <0.01; 0.01 < * <0.05).*

	Regression coefficients	**Std. error**	Dof	χ^2	**Pr(>χ^2)**
Intercept	8.2	1.1	1	96.06	2.0^{-16} ***
DepthID	-1.02	0.40	1	7.22	0.0072 *
Level	-0.87	0.41	1	4.81	0.028 *
DepthID:Level	0.093	0.16	1	0.35	0.56

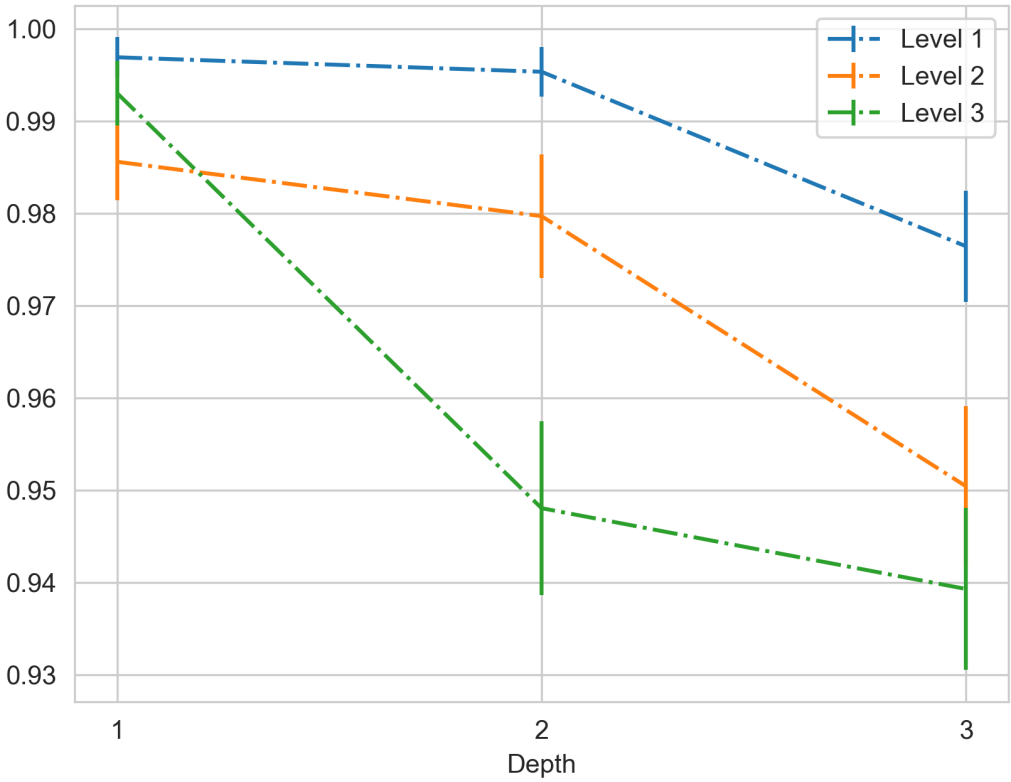

Figure 2. Success probability as a function of the 3 levels of the experiment and of the 3 depthIDs. *The figure permits appreciating that participants' performance decereases as a function of levels, but remains stable across levels when the problem planning depth is 1. See the main text for details. Error bars have been assigned according to the standard error of the mean of a binomial distribution.*

We also used a linear mixed-effect model to fit the relation between the completion time and the experimental design, with DepthID, Level and interaction as fixed effects and the subject identity as random effect. The coefficients of the regression are shown in Table 3, with their standard errors and their significance. The inclusion of the random effect was shown significant by a model comparison of the full model with and without it ($\chi^2(1,5697) = 951$, p-value < $2.1*10^{-16}$). In an analogous way, we assessed the significance of the terms by a comparison of the full model with reduced versions.

Table 3. Completion time regression. *A linear mixed-effect model is used to fit completion time. Significance code (0< *** <0.001; 0.001** <0.01; 0.01 < * <0.05).*

	Regression coefficients	**Std. error**	**DoF**	χ^2	**Pr(>χ^2)**
Intercept	4.7	1.3	1	13.41	$2.5*10^{-4}$ ***
DepthID	0.91	0.50	1	3.35	0.07
Level	2.79	0.51	1	29.87	$4.6*20^{-8}$ ***
DepthID:Level	0.85	0.23	1	13.54	$2.3*10^{-3}$ ***

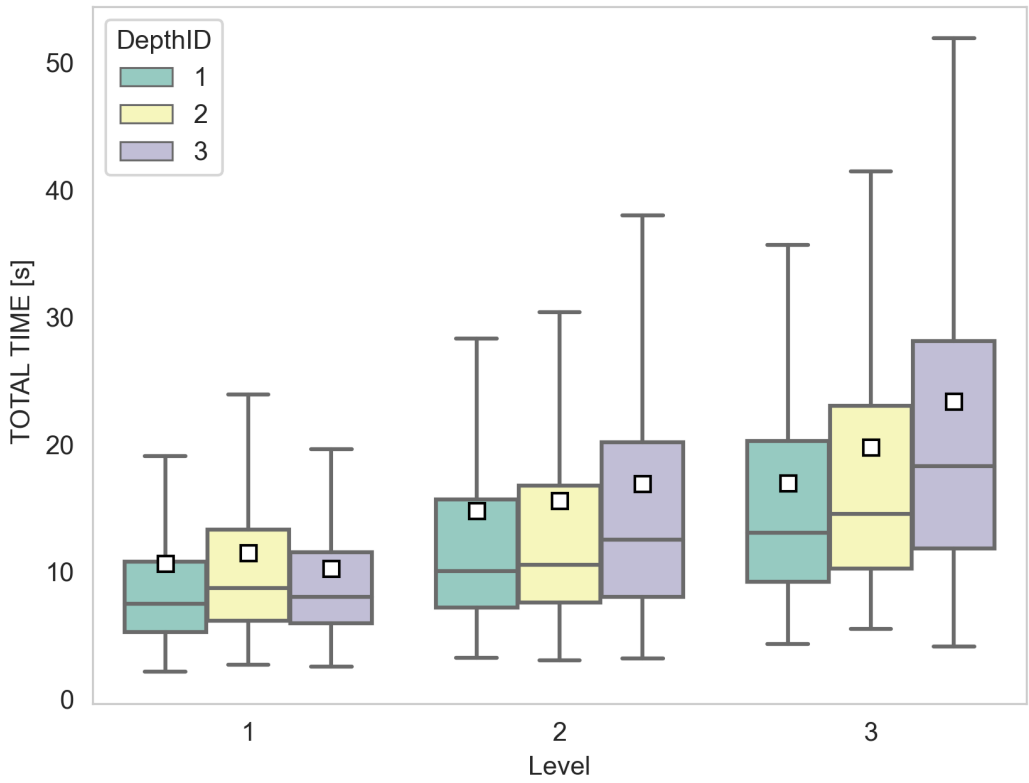

Figure 3. Problem completion time (in seconds) as a function of the 3 levels of the experiment and of the 3 planning depths. *Results are organized by level and depth. The plot shows the total time, which is bounded to 60 seconds in our experiment.*

A linear mixed-effect model for the number of backtracks (significance of the addition of the random effect: ($\chi^2(1, 5697) = 299$, p-value < $2.2*10^{-16}$) shows no main effect of DepthId and Level, but a significance for the interaction term, see Table 4.

Table 4. Number of backtrack regression. *A linear mixed-effect model is used to fit number of backtracks. Significance code (0< *** <0.001; 0.001** <0.01; 0.01 < * <0.05).*

	Regression coefficients	**Std. error**	**DoF**	χ^2	**Pr(>χ^2)**
Intercept	0.056	0.17	1	0.11	0.74
DepthID	0.014	0.072	1	0.048	0.83
Level	0.14	0.074	1	0.11	0.74
DepthID:Level	0.17	0.030	1	25.50	$4.43*10^{-7}$ ***

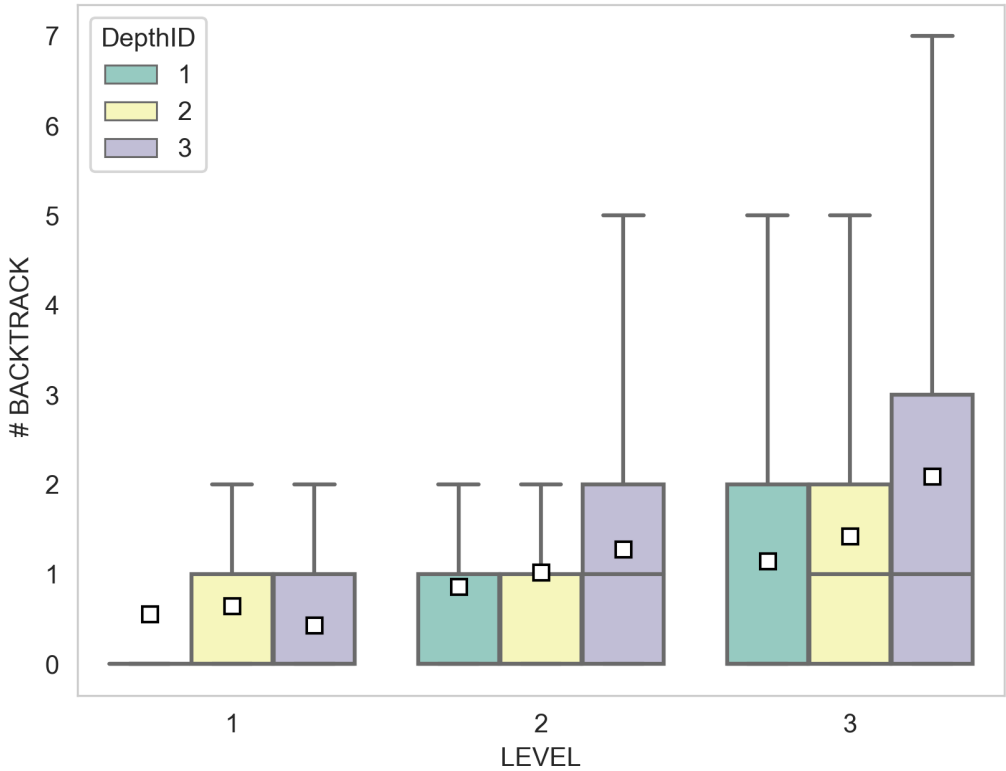

Figure 4. Total number of backtracks as a function of the 3 levels of the experiment and of the 3 planning depths. *Results are organized by level and depth.*

Participants' initial planning depth is adaptive and matches task demands

We assessed the similarity between the initial plans (i.e., the plans before the first backtrack) of participants and planning models, for each of the 8 problem groups. The results of this analysis, aggregated across the three levels and for red and blue gems, are shown in Figure 5. Our results indicate that the planner that best explains the behaviour of participants (i.e., has the maximum likelihood for the majority of problems) is the one having the minimum required planning

depth. For example, the planner that best explains the behaviour of participants during the solution of problems requiring minimum planning depth of 1 is the planner using depth 1. The same similarity between minimum planning depth of the problem and of the planner that best explains participants' data is observed across all the 8 levels, see Figure 5.

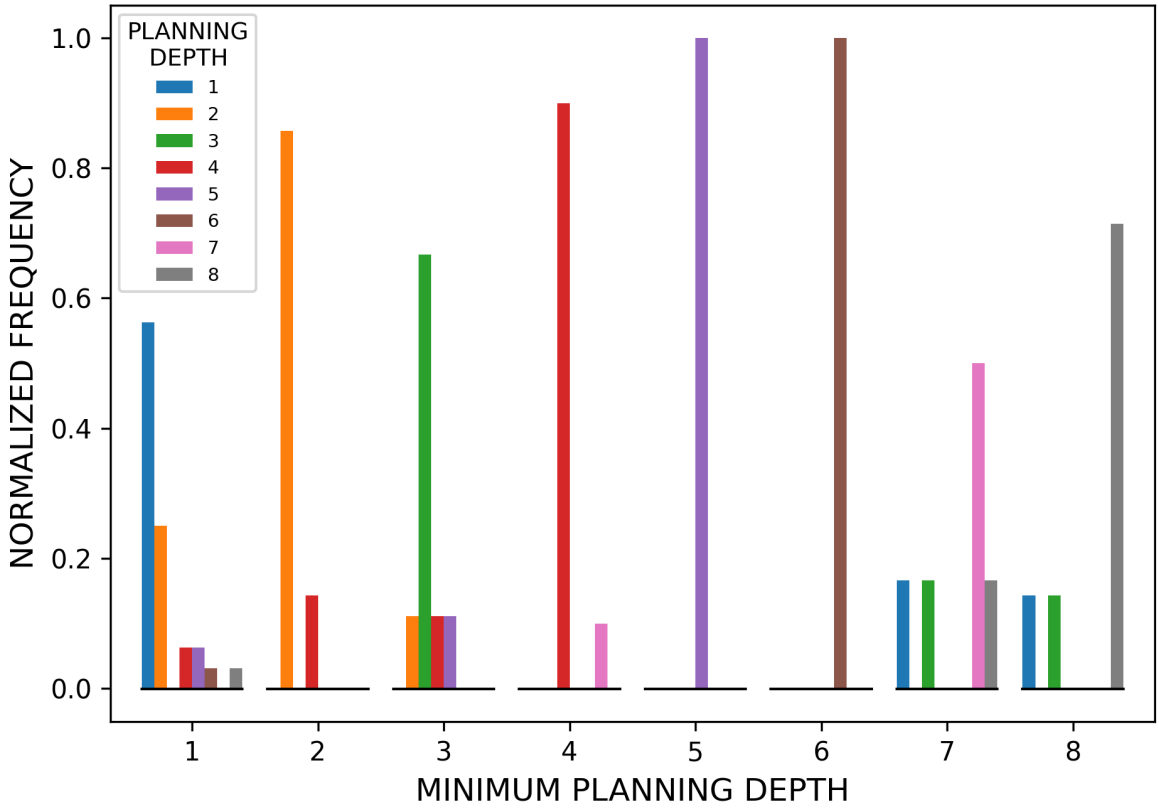

Figure 5. Comparison between participants and planning models. *The figure shows the number of times that each planner from depth 1 to 8 had the maximum likelihood of the gems collected before the first backtrack by the participants, in the problems of each of the 8 problem groups. Problems are grouped according to the minimum planning depth required to solve them, from 1 to 8, and color coded (see the legend). The frequency of each planner is normalized within each group. The figure shows that for each set of problems, from 1 to 8, the planner that wins most frequently (i.e., that has the greatest likelihood in the plurality of simulations) is the one having the minimum planning depth.*

In order to assess the statistical significance of this consideration, we tested how well using a fixed planning depth described participants' behaviour against a model where the planning depth was adapted to the minimum planning depth required by the problem. To do that, we computed for each problem the distribution of the rankings in the likelihood obtained by each possible value of the planning depth (e.g., if in a problem the model using a planning depth of 2 has the highest likelihood, it has a rank of 0 for that problem; all the other planners are ranked in a decreasing order, accordingly to the same procedure). So, we obtained a distribution of the ranks for each planning depth (90 values each). We also computed the rank of a planner that

adapts its planning depth to the minimum required for that problem. The resulting distributions are shown in Figure 6.

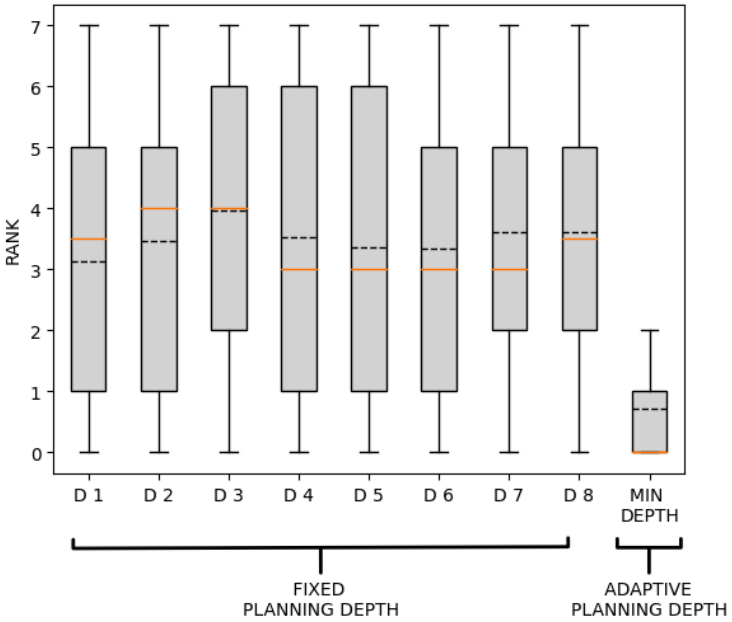

Figure 6. Rank values of fixed and adaptive depth planners across the 90 problems. *The figure shows the distributions of the ranks of planners using a fixed planning depth (D1-D8) across all problems and of the planner using a planning depth adapted to the minimum required depth to solve a problem (Min depth). The black dotted lines show mean values, whereas the orange lines show median values. See the main text for explanation.*

To test whether the average rank of the adaptive planning depth model (min depth in Figure 6) is significantly lower than each of the fixed planning depth models (D1 to D8 in Figure 6), we performed chi squared independence tests. We found that for each couple (min depth – D1, min depth – D2, etc.), the rank distributions are significantly different ($\chi^2(1, 90) = 57$, p-value $<4.0*10^{-10}$). This implies that adaptive planning depth to the minimum requirements of the problem provides the most accurate description of participants' behaviour.

We next tested for the robustness of these analyses to other methods. We obtained the same results when considering the Kullback–Leibler (KL) divergence between 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 planners (see Supplementary Figure S2). Moreover, we obtained the same results when examining the problems separately for the 3 levels of the experiment (Supplementary Figure S3), separately for red or blue gems (Supplementary Figure S4) and grouping them into classes reflecting their minimum required depth and maximum number of gems (Supplementary Figures S5 – S12). 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). Furthermore, these control analyses highlight that for all problem groups, the 8 planners collect different average gem distributions and are therefore distinguishable (Supplementary Figures S5 – S12).

As a final control simulation, we tested whether a variant of the planner used in the main analysis that knows the identity of the last gem to be collected provides a better account of participant's behavior. The motivation for control planner, which we called "obvious ending" planner, comes from the fact that some of the problems in our experiment (n = 44) had an obvious ending: if a gem was placed in a dead-end or close to it, it was necessarily the last one to be collected. We incorporated this information in the "obvious ending" planner, by calculating all the paths as in the planner used in the main analysis, but excluding all the paths where a gem in a dead end (or was only connected to dead ends) was not the last one in the plan. This procedure leads to 8 variants of the "obvious ending" planner, one for each depth level.

We then compared the performance of the 8 planners used in the main analysis and of the 8 "obvious ending" planners, by computing for all the problems how many times each model was the best fitting of participants' performance, across all problems. We found that in the problems where the two model where distinguishable, the best description was most frequently provided by the planner used in the main analysis, even though the frequency difference was minimal, see Figure S13. Finally, we found that the advantage of using an adaptive planning depth over a fixed planning depth emerges also when considering the "obvious ending" planners, showing that it is robust to information about the last gem to be collected (see Figures S14 and S15).

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. To address this question, 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. The problems were divided into three levels, characterized by increased map size and number of gems. We varied planning demands both within and between levels, therefore designing 8 groups of problems that required a minimum planning depth from 1 to 8. The planning depth required for each problem was unknown to the participants.

Our results indicate that participants' problem-solving performance depends both on planning depth and the level of the problems, underscoring that the challenge lies both in the cognitive demands imposed by a problem and its spatial scale (which increases with problem level), independently. Furthermore, the time required for problem-solving is affected by spatial scale, illustrating that addressing problems within larger maps inherently demands more time. Also the effect of the planning depth on time required for problem-solving becomes relevant, via the interaction with the spatial scale. One possible explanation is that while forming a complete plan early on is possible with smaller maps and simpler problems, working memory or attention limitations might preclude this with larger maps and more challenging problems. In other words, simpler and smaller problems would afford a global evaluation strategy, while more challenging and larger problems would require instead a more local evaluation strategy.

Additionally, the analysis of the number of backtracks shows that the spatial scale of the map and the complexity of the problem increase how often participants need to change their initial plans, but the effect size is not independent for the two dimensions.

Crucially, our results show that participants flexibly adapt their initial planning depth (i.e., the depth of their plans before the first backtrack) to the minimum depth required to solve the problems, as estimated using planning models (Figure 5). The adaptiveness of participant's planning depth is confirmed by the fact that a planner that uses the minimum depth level (from 1 to 8) for each type of problem better explains participants' behaviour, compared to a(ny) planner that uses a fixed planning depth across all problems (Figure 6). This result is robust to various control analyses: it is observed across all the 3 levels of the game (Figure S3) and the 2 types of gems, blue or red, associated to lower and higher rewards, respectively (Figure S4) and by considering planners that know the identity of the last gem to be collected, when this information is obvious to participants (Figures S13-S15).

Taken together, these results indicate that during problem solving, people make an adaptive use of their cognitive resources, by selecting an appropriate level of planning depth. Therefore, this study adds to a large literature showing that during problem solving, people adapt to the complexity and structure of the environment (Simon, 1996), make an adaptive use of their cognitive resources (Anderson, 1990) and adopt various simplifications when the problems are too challenging. 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).

An open objective for future research is assessing to what extent the results reported in this study generalize to other planning tasks. An advantage of our task is that it incentivizes planning, as opposed to (for example) moving as fast as possible to find a solution by change, because the problem space is relatively large, the subgoals (gems) are clear, there is a time incentive, and backtracking adds a time penalty. The combination of these factors might have encouraged people to plan ahead in adaptive ways, as testified by the good match with (deep) planners. At the same time, our task is not representative of all the conditions in which planning can arise, because it is static and (at least in principle) fully observable. Future studies might address situations in which certain task elements, such as edge positions or gem values, change dynamically over time. Such dynamic environments require considering how to adapt action plans to varying temporal demands and how to replan when an existing plan becomes unattainable. Similarly, future studies might address situations in which certain task elements, such as the presence or absence of a particular edge, are initially unknown or known only with a certain probability, as in the Canadian Traveller Problem (Papadimitriou & Yannakakis, 1991). Such partial observability requires balancing exploration and exploitation during planning (K. Friston et al., 2015; Sutton & Barto, 1998). Future studies might assess whether the results reported in this study generalize to dynamical and partially observable settings, or to settings that present participants with different incentives for planning versus following habits (Lancia et al., 2023).

This study has various other 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 (De Groot, 1946; Korf, 1985) and to algorithms that start from an initial "cheap" plan and progressively refine it (Krichmar & He, 2021; Lancia et al., 2023; Todorov, 2009), would impose a filter on insufficient planning depth, which could result in the over-investing bias apparent in Figure 5, when considering that for most problem groups, the second-best planner has a greater depth than the minimum required depth. 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 address the "algorithm" used by the brain to solve the problems. When establishing a similarity between the planning models 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; K. J. 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 "good enough" and 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 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 planning models do (Weld, 1994). In other words, this choice does not require delving into the full complexity of planning, acting and replanning dynamics that can occur in our setup. However, a more complete analysis of planning dynamics in our experiments should consider that backtracking is part and parcel of human problem solving (De Groot, 1946) and that people might generate longer plans when they are forced to do so – for example, because they cannot backtrack (Delaney et al., 2004). We hope to address these open issues in future studies.

Furthermore, our experiment was not designed to study whether and to what extent participants learn novel and potentially better planning strategies over time. Future studies might address learning dynamics, by comparing how people behave when presented with different sequences of planning problems – or "training curricula" (Dekker et al., 2022) – that afford or do not afford generalization across problems.

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 Agreements No. 945539 (Human Brain Project SGA3) and No. 952215 (TAILOR); the European Research Council under the Grant Agreement No. 820213 (ThinkAhead), the Italian National Recovery and Resilience Plan (NRRP), M4C2, funded by the European Union – NextGenerationEU (Project IR0000011, CUP B51E22000150006, "EBRAINS-Italy"; Project PE0000013, "FAIR"; Project PE0000006, "MNESYS"), and the PRIN PNRR P20224FESY. 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

Alterman, R. (1988). Adaptive Planning. *Cognitive Science*, *12*(3), 393–421. https://doi.org/10.1207/s15516709cog1203_3

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.

Dekker, R. B., Otto, F., & Summerfield, C. (2022). Curriculum learning for human compositional generalization. *Proceedings of the National Academy of Sciences*, *119*(41), e2205582119. https://doi.org/10.1073/pnas.2205582119

Delaney, P. F., Ericsson, K. A., & Knowles, M. E. (2004). Immediate and Sustained Effects of Planning in a Problem-Solving Task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*(6), 1219–1234. https://doi.org/10.1037/0278-7393.30.6.1219

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
Friston, K., Rigoli, F., Ognibene, D., Mathys, C., Fitzgerald, T., & Pezzulo, G. (2015). Active inference and epistemic value. *Cogn Neurosci*, *6*(4), 187–214. https://doi.org/10.1080/17588928.2015.1020053
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
Gershman, S. J., Horvitz, E. J., & Tenenbaum, J. B. (2015). Computational rationality: A converging paradigm for intelligence in brains, minds, and machines. *Science*, *349*(6245), 273–278. https://doi.org/10.1126/science.aac6076
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
Hayes-Roth, B., & Hayes-Roth, F. (1979). A cognitive model of planning. *Cognitive Science*, *3*(4), 275–310.
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*.
Krusche, M. J. F., Schulz, E., Guez, A., & Speekenbrink, M. (2018). *Adaptive planning in human search* (p. 268938). bioRxiv. https://doi.org/10.1101/268938
Kuperwajs, I., Ho, M. K., & Ma, W. J. (2024). Heuristics for meta-planning from a normative model of information search. *Planning*, *1*, a2.

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.
Papadimitriou, C. H., & Yannakakis, M. (1991). Shortest paths without a map. *Theoretical Computer Science*, *84*(1), 127–150.
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.
Simon, H. A. (1996). *The sciences of the artificial* (Vol. 136). MIT press. https://books.google.it/books?hl=en&lr=&id=k5Sr0nFw7psC&oi=fnd&pg=PR9&dq=simon+%22The+Sciences+of+the+Artificial%22&ots=-v-HpEENDC&sig=I1fjzb5jB-3qHaNYbTP7l7AIADw
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
Solway, A., Diuk, C., Córdova, N., Yee, D., Barto, A. G., Niv, Y., & Botvinick, M. M. (2014). Optimal Behavioral Hierarchy. *PLOS Computational Biology*, *10*(8), e1003779. https://doi.org/10.1371/journal.pcbi.1003779
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., in ‘t Veld, L., & Hennink, T. (2021). Motivation as a function of success frequency. *Motivation and Emotion*, *45*, 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., Kuperwajs, I., Galbiati, G., Bnaya, Z., Li, Y., & Ma, W. J. (2023). Expertise increases planning depth in human gameplay. *Nature*, 1–6. https://doi.org/10.1038/s41586-023-06124-2
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.
Wilensky, R. (1981). Meta-planning: Representing and using knowledge about planning in problem solving and natural language understanding. *Cognitive Science*, *5*(3), 197–233.