SAPIENZA
UNIVERSITÀ DI ROMA

Exploring Human Planning Strategies in the ThinkAhead Task: Combining Computational, Kinematic, and Eye Movement Approaches

Scuola di dottorato in Scienze Psicologiche
PhD in Psychology and Cognitive Science (XXXVI cycle)

Mattia Eluchans
ID number 1619026

Advisor
Dr. Giovanni Pezzulo

Co-Advisor
Prof. Stefano Sdoia

Academic Year 2023/2024

Thesis defended on 17/09/2024
in front of a Board of Examiners composed by:

Prof. Nicola Cellini, Raffaella Nori, Prof. Anna Borghi (chairman)

Exploring Human Planning Strategies in the ThinkAhead Task: Combining Computational, Kinematic, and Eye Movement Approaches
PhD thesis. Sapienza University of Rome

This thesis has been typeset by LaTeX and the Sapthesis class.

Author's email: mattia.eluchans@uniroma1.it

Abstract

Planning is a ubiquitous process in everyday life, but its covert nature makes it difficult to study, and consequently, various paradigms and methods have been proposed to do it.

In this thesis, we investigated human planning strategies in a novel task inspired by the Travelling Salesman Problem (TSP), that we called the ThinkAhead task, utilizing a combination of computational, behavioral, kinematic and eye-tracking measures. The task required participants to solve a sequence of problems by finding a path that connects a set of nodes placed in a 2D grid-like graph without passing twice over the same node (*simplicity condition*). The task's multistep nature, combined with the simplicity requirement and a time limit, made it challenging enough to require planning. We collected data about this task in two different setups: at first, online, where participants completed the task on their mobile devices; then, in-person, where we also recorded gaze and mouse movements.
We then attempted a characterization of their planning strategies from different perspectives.

First, we employed a computational approach, and compared the behavior of participants with that of artificial agents whose planning depth was manipulated. We then tested whether the use of cognitive resources in this modeling paradigm was better explained by a fixed planning depth or a flexible one that adapted to the task demands. The results showed that the latter adaptive strategy provided a better fit to the participants' behavior.

Secondly, we looked for signatures of planning in the kinematics of mouse trajectories and gaze behavior. We focused on instances of fluid movements, where no pauses occurred, and observed an alternating pattern in eye-hand coordination. This pattern involved a gaze fixating a target followed by a reaching movement towards the fixation coordinates. Within this descriptive paradigm of motor plans sequences, we searched for evidence of coarticulation, which is generally defined as a phenomenon in which the execution of subtasks is influenced by the overall task: for example sub-movements are different if part of the same motor plan rather than independent.
To formally test this, we trained a (linear) classifier for the next action on: (1) segments of mouse trajectories, and (2) gaze position on the screen. In both cases, even though with different performance, the classifier's accuracy was significantly higher than chance level, which means that current actions carry information about future one. This coarticulation provides evidence for planning in this task. Additionally,

we tested whether gaze behavior during a fixation could predict the direction of the next fixation above chance level and beyond the predictions of models based purely visual information.

Finally, we studied the commitment of participants to their plans, and the possible determinants of backtracking. We found that backtracking was in most of the cases performed only after a visual feedback of unsolvability of the task (for example because the map had been bisecated and there were rewards to be collected on both the (dis)connected components). This suggests that participants did not perform a continual (or highly frequent) replanning, but rather relied upon perceptual messages of errors. We then proposed a task-specific taxonomy that might help differentiate the signals that triggered the backtracking event.

Contents

1 Introduction **1**

1.1 Planning: a fundamental cognitive process 1

1.1.1 Planning as a search process 1

1.1.2 Heuristics 3

1.1.3 Encoding and Hierarchical planning 3

1.2 Studying planning: how to probe a covert process 4

1.2.1 Computational Model-Based Analysis 5

1.2.2 Continous Movement Kinematic Recordings 6

1.2.3 Eye tracking 7

1.3 Structure of the thesis 9

2 ThinkAhead: a novel task to study human planning **11**

2.1 Planning Depth 12

2.2 Experiment design 14

2.3 Discussion 19

3 Adapting planning depth in problem solving **21**

3.1 Introduction 21

3.2 Methods 23

3.2.1 Data collection 23

3.2.2 Analysis of success probability, problem completion time and total number of backtracks 23

3.2.3 Similarity between participants and planning models, for each of the 8 problem groups 24

3.3 Results 24

3.3.1 Participants are sensitive to both problem level and planning depth, but in different ways 24

3.3.2 Participants' initial planning depth is adaptive and matches task demands . . . 29
3.4 Discussion . . . 31

4 Coarticulation in eye-hand coordination as a signature of planning 37
4.1 Introduction . . . 37
4.2 Methods . . . 43
4.3 Results . . . 47
4.3.1 Eye-Hand Alternated Dynamic . . . 47
4.3.2 Spatio-Temporal Principal Component Analysis (st-PCA) reveals meaningful structure in the kinematics . . . 51
4.3.3 Detecting Coarticulation in the Hand Kinematics and Eye Position . . . 53
4.3.4 Predicting next fixation: coarticulation in the eye angular position . . . 59
4.4 Discussion . . . 62

5 The Determinants of Backtracking: Evidence against continual replanning 65
5.1 Introduction . . . 65
5.2 Methods . . . 66
5.2.1 Participants and Experimental Setup . . . 66
5.2.2 Determining the task solvability . . . 66
5.2.3 Creating a taxonomy of backtracking events . . . 67
5.3 Results . . . 69
5.3.1 Continual monitoring of a plan or visual feedback? . . . 69
5.3.2 The determinants of backtracking: a taxonomy . . . 71
5.4 Discussion . . . 72

6 Conclusions 75

A Supplementary materials 81

Bibliography 87

Chapter 1

Introduction

1.1 Planning: a fundamental cognitive process

Thinking about the future is a frequent attitude in our life. At different timescales and in different contexts, we are constantly imagining possible futures, and planning our actions accordingly: we think and make plans about what to do at the current moment, day, week or further periods of time. We plan our behaviour even out of time, and decide how we would behave in a hypothetical situation, and we do this in many different tasks, such as deciding what to eat, how to commute, how to win current chess game etc.

Despite the intrinsic heterogeneity and richness of these tasks, the initial attempts to explain planning shared the belief that a reductionist and formal approach could uncover a general human strategy. Planning was assimilated to *problem solving* (B. Hayes-Roth & Hayes-Roth, 1979a; McDermott, 1978; Newell, Simon, et al., 1972; Sacerdoti, 1975), and consequently a *planner* was essentially as a *problem solver*.

1.1.1 Planning as a search process

As stated in (McDermott, 1978):

> A problem solver operates by reducing problems to simpler subproblems repeatedly until immediately solvable (primitive) problems are reached. The result is a plan which must be executed in order to realize a solution to the problem.

This procedural description of problem solving — and thus of planning — formulates the hypothesis that the planner can identify subgoals and decompose

the task into simpler components. Planning is then conceptualized as the proccess of search across this simpler elements of the problems, and of the best course of action to achieve a goal (Craik, 1943; Korf, 1987; Sacerdoti, 1974), performed over a mental map of the task (Tolman, 1948). This planning formalization is still valid nowadays (Daw & Dayan, 2014; Hunt et al., 2021; Kolling et al., 2012). If we look for example at the popular paradigm of reinforcement learning (Sutton & Barto, 1998), a planner has a representation of the value of the states of the world, and chooses the best action to do only after having estimated the value of the possible future trajectories.
Similarly, the so-called paradigm of *planning as inference* (Botvinick & Toussaint, 2012; K. Friston et al., 2015; K. J. Friston et al., 2016) frames this search in a slightly different way: a planner has a probabilistic representation of the world and its dynamics and knows its goal, and tries to infer the course of actions that will lead to it with the highest probability.

Because of this idea of a search among different possible futures, several planning studies in humans and other animals utilized tree-like structure for tasks, primarily focusing on simple (e.g., two-step) problems that could be exhaustively searched (Akam et al., 2015; Daw et al., 2011; Hasz & Redish, 2018; Miller et al., 2017). Organizing the task as a tree allows for a clear representation of the possible actions and outcomes, and the planner can then search through the tree to find the best course of action.

Depending on the task to be solved, the combinatorial nature of the planning can make an exhaustive search an infeasible strategy to be followed, especially if time or cognitive resources are limited. This requires planners to trade off the accuracy of the solution with their complexity. In the game of chess, for example, the number of possible moves increases exponentially with the number of steps ahead that the player wants to consider (Shannon, 1988). In the same way, while commuting, the number of possible routes increases with the number of streets that the driver can take, and there are many different factors that one can include or exclude from this computation. There are many factors that have an impact in the complexity of the exploration: the state space, the space of the actions, the space of goals, the possibly stochastic interaction with the environment, and so on. Furthermore, planning exhaustively, even if possible, might be costly in terms of time and cognitive resources and an exhaustive search might not be worth the effort (Russek et al., 2022).

This is a possible instance of the well-known problem of the exploration-exploitation dilemma: when is it worth keeping on exploring the possible plans and when is it better to exploit the current knowledge to make a decision (K. Friston

et al., 2015; Sutton & Barto, 1998). Different models consider different principles to address this problem.

1.1.2 Heuristics

Once assumed that evaluating *everything* is too hard, the problem of establishing *how much* and *what* to evaluate emerges (Wiener et al., 2009).

Among the diverse possible strategies that reduce the cognitive cost of planning, heuristics represent a frequently used approach in search problems (Geffner & Bonet, 2013; Russell & Norvig, 1995). A heuristic is an approximate strategy that intend to solve a task without performing all the necessary computations to be optimal. It is a - possibly biased - strategy that simplifies the problem at hand to make it more tractable. The origins of these heuristics might be a form of adaptation of thinking process to the task properties (Anderson, 1990; Simon, 1996), derive from a rational usage of the resources available to the planner (Bhui et al., 2021; Callaway et al., 2018, 2022; Lieder & Griffiths, 2020; Simon, 1957), be induced by the time pressure (Van Opheusden et al., 2017) or by some underlying objective function (Solway et al., 2014; Tomov et al., 2020). In (Q. J. Huys et al., 2012) - and also (Balaguer et al., 2016) - it is proposed that people prune unpromising branches of the search tree, which means implicitly assuming that nothing "too good" will happen in the direction of a specific branch, for example because the firsts "steps" in that directions where extremely "negative".

A single heuristic is usually not enough to encompass all the complexity of human behaviour even in simple tasks. For this reason it can also be beneficial to explain it by considering an interplay of many different strategies (Q. J. M. Huys et al., 2015). Other heuristics could be sampling only few promising routes, or many routes but only up to a certain depth (Keramati et al., 2016; Pezzulo et al., 2013).

In the widely-used tree search description this resource rational usage might mean 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; B. Hayes-Roth & Hayes-Roth, 1979b; Kuperwajs et al., 2024; Wilensky, 1981).

1.1.3 Encoding and Hierarchical planning

A second stream in the planning literature mitigates cognitive costs exploits the *encoding* phase that preceeds the planning phase. Planners form a mental represen-

tation of the problem upon which they base the mental search. Previous studies suggest that during the encoding phase, participants might form simplified mental representations of the problems, omitting many details of the original task (Ho et al., 2022). The idea of using a cognitive map (Tolman, 1948) to redefine the state space in which to plan can be expressed in many ways: it could be grounded in information properties of the map and its possible transitions (thus, augmenting the simple notion of topology) in (Maisto et al., 2016; van Dijk & Polani, 2011). This compression can make the planning easier by reducing the number of states that need to be considered, or by exploiting map properties that are useful to the task (bottlenecks, isolated states etc.).

This encoding phase that creates a higher-level map of the problem, naturally leads to hierarchical models for planning (Balaguer et al., 2016; Pezzulo et al., 2018; Ribas-Fernandes et al., 2011; Solway et al., 2014). In such models, the planner can decompose the task into more manageable subgoals, and then dynamically switches between these two different levels of abstraction of the task, partly planning at a higher-level and thanks to that getting constraint for the planning in the original space. There are some general benefits from this approach (Maisto et al., 2015), and as demonstrated in (Donnarumma et al., 2016) this could also explain aspects of human performance in the Tower of Hanoi task. By leveraging multiple levels of abstraction, the planner can optimize its strategy, balancing the need for detailed planning in complex situations with the efficiency of higher-level strategies for more straightforward scenarios.

This hierarchical approach not only mitigates the computational demands of planning but also reflects a more realistic model of human cognitive processes, where individuals often break down complex problems into simpler components and adjust their level of detail based on the task at hand, similar to the general problem solver previously quoted. Reaching a distant location when in a big city is a concrete task where these hierarchical strategies are useful: one can first plan some intermediate steps (for example reaching a metro station) and then plan the details of the path to reach it.

1.2 Studying planning: how to probe a covert process

Both in past (Craik, 1943; Shallice, 1982) and modern descriptions (Daw & Dayan, 2014; Daw et al., 2011; Hunt et al., 2021) of planning, the formalization of cognitive processes as computational processes has been a common approach. However, one

of the main challenges in studying planning is that it is a covert process occurring within the brain, making it difficult to observe and measure directly.

The covert nature of planning means that researchers cannot directly observe the cognitive steps individuals take when formulating plans. Instead, they must rely on indirect measures to infer these processes. This necessitates the creation of experimental setups that capture behavioral or physiological outputs somehow related to the underlying planning mechanisms. For instance, researchers may use eye-tracking to monitor gaze patterns, analyze mouse movements to infer decision-making processes, or employ neuroimaging techniques to observe brain activity during planning tasks. These methods help to bridge the gap between the unobservable cognitive processes and the measurable external behaviors, allowing for a deeper understanding of how planning operates within the brain.

By combining data from these various indirect measures, researchers can build robust models of the planning process, elucidate the cognitive architecture underlying planning, and also inform the development of artificial intelligence systems that emulate human planning strategies.

1.2.1 Computational Model-Based Analysis

One viable approach to study planning is to directly look at the outputs of a task, namely the "choices" made by the subject. The state space in which these choices exist is usually task-specific, and range from a move in a game (Chase & Simon, 1973; De Groot, 1946; Krusche et al., 2018), to a route on a map (Lancia et al., 2023; McDermott & Davis, 1983; Wiener et al., 2009).

To provide possible explanations for this behavior, one could develop artificial model-based agents that make a "choice" within the same state space as the human. Thus, modeling involves designing algorithms that can solve a task and then comparing the models' solutions to those of humans. This comparison can occur at various levels of abstraction. For example, one could assess whether the choices made by the model match those made by the human, or whether the model can predict human choices. Alternatively, one might compare specific properties of the choices made by the model with those made by the human, such as the time taken to make the choice, as in the drift-diffusion model (Fudenberg et al., 2019; Myers et al., 2022; Ratcliff, 1978; Stone, 1960).
This approach has interesting implications when the parameters of the model are fitted to the participants, and correlated with other traits (even psychiatric), allowing for a form of computational phenotyping (Q. Huys, 2014; Q. Huys et al., 2020; Kafadar et al., 2022; Montague et al., 2012).

Since fitting a model does not retain absolute meaning, it is generally necessary to perform a model comparison on the same benchmark (MacKay, 2003). A fundamental and widely used tool is the comparison of the likelihood function. This function evaluates the fit of a model to the observed data and quantifies the probability of the data given the model's parameters, serving as a measure of model adequacy. Models can be compared using likelihood-based criteria, such as the Maximum Likelihood, the Akaike Information Criterion (AIC), or the Bayesian Information Criterion (BIC), which incorporate both the likelihood of the model and the complexity of the model in terms of the number of parameters. This approach enables researchers to select models that most adequately describe the data while avoiding overfitting, ensuring that the model generalizes well to new data.

1.2.2 Continous Movement Kinematic Recordings

A choice in the decision making task represents just the final output of a reasoning process but the underlying process and mental search remain unexpressed. The analysis of response time, i.e. the time taken before making a choice, as done in (Fudenberg et al., 2019; Myers et al., 2022; Ratcliff, 1978; Stone, 1960) already goes beyond the simplest output, as it can be related to the computation needed to make that choice, and how "hard" making that choice is. However, in this description, action and decision are assumed to be performed in a purely sequential way: evidence for a choice is accumulated and *after* that, a choice is made. However, decisions in real life can't always be made in this way. As stated in the paper: (Maselli et al., 2023):

> [...] soccer players have to acquire information about the state of the game while running and deciding where to position themselves, or whether to pass or shoot the ball. Similarly, a car driver has to monitor the road and the movements of the other cars while continuously making decisions about where to turn, whether to accelerate or stop the car, etc.

In these cases the decision process is not separated from the action process, but rather they are intertwined. The decision and the acting processes can be performed in parallel, or they can influence reciprocally.

The kinematic analysis of the movements that implement a choice can provide a window into the decision process (Barca & Pezzulo, 2012; Freeman et al., 2011; Song & Nakayama, 2009). In (Barca & Pezzulo, 2012) authors considered a binary discrimination task where the selection of the correct category was done via mouse selection. They found that increasing the ambiguity of the stimuli increased the

mouse trajectory curvature. This was interpreted as a sign of the decision process unfolding over time, compatible with a paradigm of parallel processing of action and decision. This sort of information could not be obtained by just looking at the final choice, such in tasks where the choice is a simple button press.

This dependence of the modalities of execution on the underlying planning process is obviously not limited to the simplified domain of the mouse movements. In (Maselli et al., 2017) the analysis of the *modalities* of the implementation of a choice, rather than the choice itself, is made explicit: in a throwing motor task the predictability of the direction of the throw was analyzed by looking at the kinematics of the whole body movement. Also in (Maselli et al., 2019) throwing styles were analyzed by looking at the whole-body kinematics and providing a characterization much richer than the simple choice that rested upon the coordination of different body parts.
We could say that in general, the kinematic domain expands significantly the possibility of analysis of a decision process and represents an interesting factor to be considered in the design of an experiment (Krotov et al., 2022).

1.2.3 Eye tracking

Since the pioneering works of Yarbus (Yarbus, 1967) the importance of the eye tracking in the study of cognitive processes, along with its complexity and richness, has been recognized. A commonly accepted, and purely kinematic description of eye movements, typically identifies few main different regimes of movements:

- Fixations. These are periods when the eyes remain stationary and focused on a single point.
- Saccades. These are rapid, (mostly) ballistic movements that shift the eyes from one point to another.
- Smooth pursuit. These are slow, continuous movements that track a moving object.

This minimalistic description of eye movements is based upon the temporal dynamic, but why a fixation is made on a certain point, or why a saccade is made in a certain direction and at a certain time, is a much more complex question that rests upon *both* sensory *and* cognitive processes.

Eye movements can in fact - and unsurprisingly - be related to the characteristic of a visual input. In (Mannan et al., 1997) similarities in fixation patterns among

subjects observing the same stimuli were found to have a correlation with the edge density of the input image. Also in (Itti & Koch, 2001; D. Parkhurst et al., 2002; D. J. Parkhurst & Niebur, 2003) the importance of the saliency, intended as how much a region of the image stands out from the background, was recognized as a determinant of the fixation position.
However, next to these bottom-up direction, the determinants of eye movements can also be the end of a top-down stream. It was already shown in (Yarbus, 1967) that the gaze position was influenced by what was the goal of the observer (or almost dually: by the instructions given by the experimenter).

This task-dependency becomes even more interesting when some sort of control or interaction with objects is considered. In many of Land works (M. Land, 2004; M. Land & McLeod, 2000; M. F. Land et al., 1999) naturalistic tasks (making tea, playing cricket etc.) show that the eyes fixate regions of interest for the specific task and are strongly coupled with the phase of the task (e.g. the different moments composing the making of tea). These considerations can lead to the augmentation of the classification of eye movements beyond the simple kinematic as in (M. F. Land et al., 1999) where a richer taxonomy of fixations based upon their function (such as "monitoring", "planning" etc.) was used. Also, the very same properties of these eye movements can be task-dependent. If we look at the duration of the fixations, it is usually assumed that it ranges between 200 ms and 400 ms, but for example looking at the widely-studied task of reading, it is reported in (Radach et al., 1999) that there can be shorter fixations which last only 50 ms.

Fixations can also be related to the planning of future actions. In (Pelz & Canosa, 2001) found that in a complex task most of the fixations were directed towards objects of immediate use, but a smaller number of times it was directed towards objects whose use would be needed in the future. An even richer analysis of the temporal dynamics of gaze behaviour (Zhu et al., 2022) shows a consistent presence of backward and forward planning during a route planning task. In (Huang et al., 2023) a variation of the tic-tac-toe game is implemented and played while recording gaze behaviour. It is found that participants gaze showed a predictive ability with regard to the next moves, which increased with participants' expertise.

Similarly to fixations also saccades can be related to the planning of future actions. In (De Vries et al., 2014) the authors found evidence of planning during scanpath in a 2-saccades switching task where a modification of a planned sequence was induced by a change in the visual stimuli. Also in (Baldauf & Deubel, 2008) the authors performed a discrimination task to study attention allocation during sequences of saccades and found that there is evidence that rather than being planned purely sequentially, they are planned simultaneously, similarly as done

in (McPeek et al., 2000) or (Ames et al., 2019). In (Azadi et al., 2021) they found that the saccade amplitude and direction was modulated by the future saccade in a sequence.

Many of these research have been made possible by the advancement in eye tracking technology, from the mirror connected to the eye in (Yarbus, 1967) or (Robinson, 1963). A less invasive and more powerful method of eye tracking is provided by optical infrared-based systems. These systems are based on the principle of the reflection of infrared light from the cornea and the pupil of the eye: the corneal reflection is invariant to the eye movement, while the pupil reflection moves with the eye. The difference between the two reflections is used to calculate the gaze direction.

The other significant advancement in eye-tracking technology is the development of portable devices. This advancement has allowed researchers to study gaze behavior in more naturalistic settings, moving beyond the constraints of laboratory environments. For instance, studies by (M. F. Land et al., 1999) and (M. Land, 2004) utilized portable eye-tracking systems to observe naturalistic gaze behavior, providing valuable insights into how eye movements and planning processes function in real-world scenarios. These portable devices have thus expanded the scope and ecological validity of eye-tracking research, enabling the exploration of planning and attention in more diverse and realistic contexts.

1.3 Structure of the thesis

The rest of the thesis is structured as follows.

Chapter 2

We introduce the task developed and used to study planning throughout this work: the ThinkAhead task. We describe the task and the main features that make it a good candidate to study planning.

Chapter 3

The main results of the first (online) study are presented. We explain how we used a mode-based approach to compare the behaviour of participants with the behaviour of artificial agents whose cognitive resources were manipulated. We then tested if the usage of the cognitive resources in this modelling paradigm was better

explained by the usage of a constant planning depth or by the usage of flexible planning depth which adapted to the task demands. Results showed that indeed the latter adaptive strategy was the one that better explained the behaviour of participants.

Chapter 4

The results from the second experiment, where the task was performed in-person while eyes and hand movements were being tracked, are presented. We looked for signature of planning in the kinematics of the mouse trajectory and the gaze behaviour. In particular, we considered the case of fluid movements, where no pauses or backtracks occourred, and noticed an alternating behaviour in the eye-hand coordination, that consisted in a gaze fixation on the target (node) followed by a reaching movement towards it. This alternation provided us with a chunking paradigm for the path implementation, and provided with a useful setup to test if the kinematic behaviour of the mouse and of the fixation position were influenced by the future actions. To do this, we used a linear classifier to predict the next action from the mouse trajectory segments and gaze positions on the screen. Both methods, despite differing in performance, significantly outperformed chance, suggesting that current actions provide information about future ones via the modalities in which it has been performed. This supports the presence of planning in the task. We also tested the relation between gaze behavior during a fixation and the direction of the next fixation better than chance and more accurately than models based on purely visual data (and no planning).

Chapter 5

In the last chapter we explore the commitment of participants to their plans. By looking at the moments where participants backtracked during the ThinkAhead task, we found that in most of the cases there were strong visual feedback of non-solvability of the task. This could suggest that participants did not always predict the consequences of their actions before the movement was initiated, and rather made a move and looked for a feedback signal, as in off-loading strategies. We then proposed a taxonomy that could differentiate the signals that trigger a backtracking event.

Chapter 2

ThinkAhead: a novel task to study human planning

As stated in the introduction, one of the striking features of planning is its presence across many contexts. Some of these are more explicitly linked to planning processes, as in chess (Chase & Simon, 1973; De Groot, 1946; Gobet, 1998; Krusche et al., 2018; Saariluoma, 1995), the Travelling Salesman (MacGregor & Ormerod, 1996; Vickers et al., 2001), the Canadian Traveler Problem (Papadimitriou & Yannakakis, 1991) or the Tower of Hanoi (Donnarumma et al., 2016; Maisto et al., 2015; Shallice, 1982). Each of these tasks highlight different aspects of the planning process, and have been used to study different aspects of it.

In this chapter, we introduce a novel task, the ThinkAhead task, that we designed to study planning in a controlled and systematic way. The task is inspired by the Travelling Salesman Problem (TSP) and consists on a series of problems that request finding a path in a 2D grid-like graph to connect all the "gems" in a grid, without passing twice over the same node. However, nodes can be unselected by backtracking.

In our setups participants solved the problems by "navigating" with the finger (touch control) or a cursor in a grid that was fully visible on a screen.
Figure 2.1 shows an example problem, which requires finding a path from the start location (yellow node) through all the gems (red nodes). The six panels show six representative timesteps of the solution, with the links colored in blue indicating the selected path on the graph (visible to participants), the azure points showing the past trajectory (invisible to participants) and the small red dot showing the actual gaze position at different times (invisible to participants). In the example illustrated in Figure 2.1, the participants first see on the screen 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.

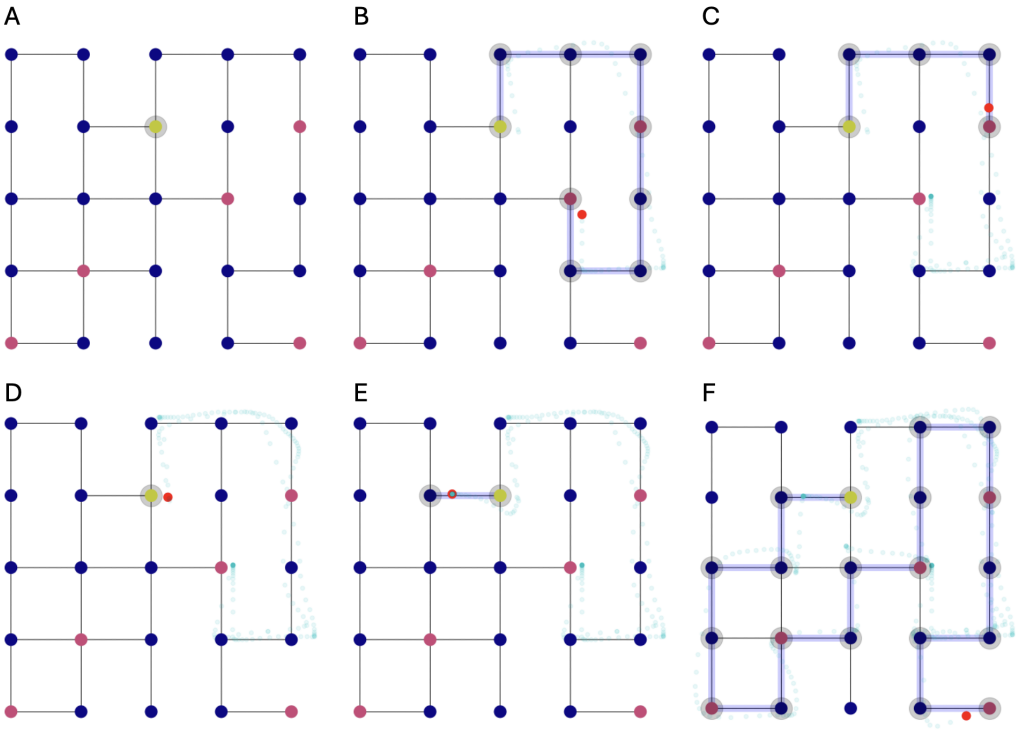

Figure 2.1. An example problem in the experiment. The problem requires finding a path in the grid that starts from the home location (yellow node) and collects all the "gems" (red nodes), without passing through the same node twice. Participants solved the problem by navigating with their finger on the (fully visible) grid. 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.

This example highlights the richness in data that are provided by this task. As will be explained in next section 2.1 problems that require different planning horizons can be easily constructed, and the task can be used to study different aspects of planning, such as the planning depth, the commitment to a plan, the determinants of backtracking, and the kinematics of the movements.

2.1 Planning Depth

The simplicity condition, i.e. the fact that nodes cannot be crossed twice, and the multiple objectives makes (a certain amount of) planning necessary to solve most

of the problems. Crucially, we designed problems that require different planning depths to be solved.

The concept of planning depth has not a unique definition: in general it is 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), while in other problems 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 2.1 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 the 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.

Assessment of the minimum planning depth required for each problem

To assess problem minimum required planning depth 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 considered 8 variants of the same planner, which only differ for the depth parameter (n), ranging 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 among them 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 (i.e. no more rewards can be collected from that position). 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.

We varied the planning depth parameter 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.

2.2 Experiment design

The experiment comprised 90 problems. 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, 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 2.1. 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 2.1. 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 2.1. 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 = 2-4 Size = 6x8 Gems = 6-7
Sublevel 2	Depth = 2-4 Size = 5x5 Gems = 7-8	Depth = 2 Size = 6x6 Gems = 7-8	Depth = 5-8 Size = 6x8 Gems = 4-5
Sublevel 3	Depth = 5-8 Size = 5x5 Gems = 4-5	Depth = 3 Size = 6x6 Gems = 6-7	Depth = 5-8 Size = 6x8 Gems = 7-8

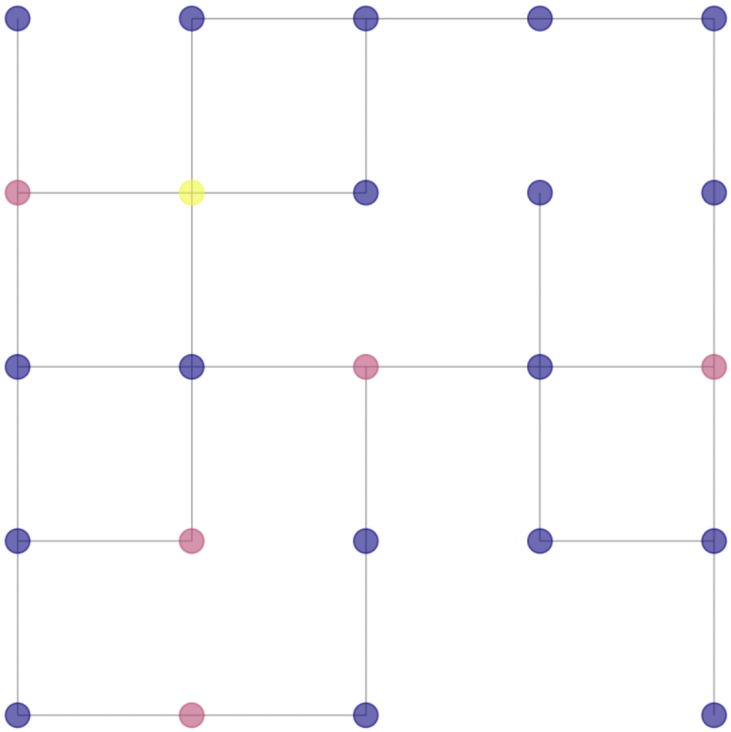

Figure 2.2. An example problem from the first experiment. The problem requires finding a path in the grid that starts from the home location (yellow node) and collects all the "gems" (red nodes), without passing through the same node twice.

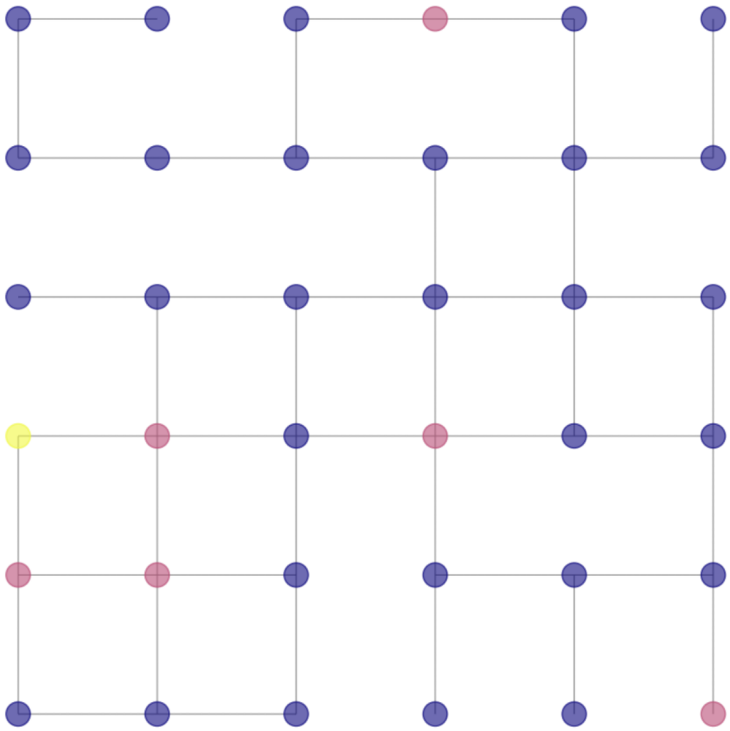

Figure 2.3. An example problem from the second experiment.

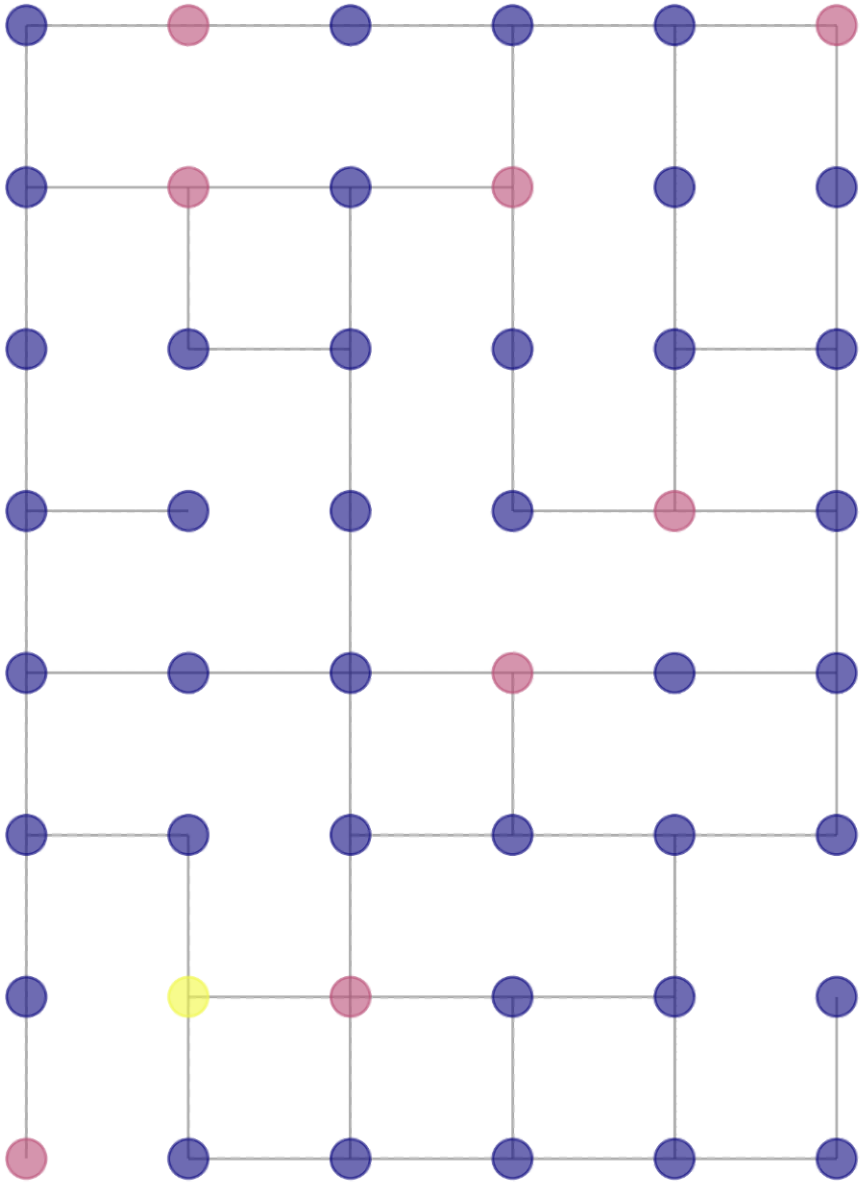

Figure 2.4. An example problem from the third experiment.

2.3 Discussion

The ThinkAhead task offers a rich platform for analyzing various dimensions of human planning and problem-solving. Through this task, we can combine standard behavioral measures analysis, such as success rates and completion times, with model-based analyses of planning strategies. Additionally, the participants' path implementation via touchscreen or cursor continuous movement, along with the integration of eye tracking, can provide deep insights into the dynamic interplay between visual attention and motor actions, revealing patterns of coarticulation and continuous planning.

Chapter 3

Adapting planning depth in problem solving

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 compared the performance of a group of human participants with some artificial planning models in the ThinkAhead task. We found that 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.

3.1 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, et al., 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 (Q. J. M. Huys et al., 2015, 2012; 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 et al., 2014; 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 evaluating 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 et al., 2023) 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; B. Hayes-Roth & Hayes-Roth, 1979b; 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.

3.2 Methods

3.2.1 Data collection

The experiment was conducted with the support of ThinkAhead, an Android application developed to study navigational planning and problem solving. 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 ± 10years) who tried at least 80 of the 90 problems of the experiment, although they did not necessarily solve all of them.

3.2.2 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 analyze participants' success probability (i.e., the probability that they solved the problems before the deadline), 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).

3.2.3 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 possible values of the planning depth parameter). 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.

3.3 Results

3.3.1 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 3.1, 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: ($\chi^2(1, 5697) = 112, p - value < 2.1 * 10 - 16$).

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 3.2, 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

Table 3.1. 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	Chi2	Pr(>Chi2)
Intercept	8.2	1.1	1	96.06	2.0e-16 ***
DepthID	-1.02	0.40	1	7.22	0.0072 *
Level	-0.87	0.41	1	4.81	0.028 *
Interaction	0.093	0.16	1	0.35	0.56

full model with reduced versions.

Table 3.2. 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	Chi2	Pr(>Chi2)
Intercept	4.7	1.3	1	13.41	2.5e-4 ***
DepthID	0.91	0.50	1	3.35	0.07
Level	2.79	0.51	1	29.87	4.6e-8 ***
Interaction	0.85	0.23	1	13.54	2.3e-3 ***

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.

Table 3.3. 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	Chi2	Pr(>Chi2)
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
Interaction	0.17	0.030	1	25.50	4.43e-7 ***

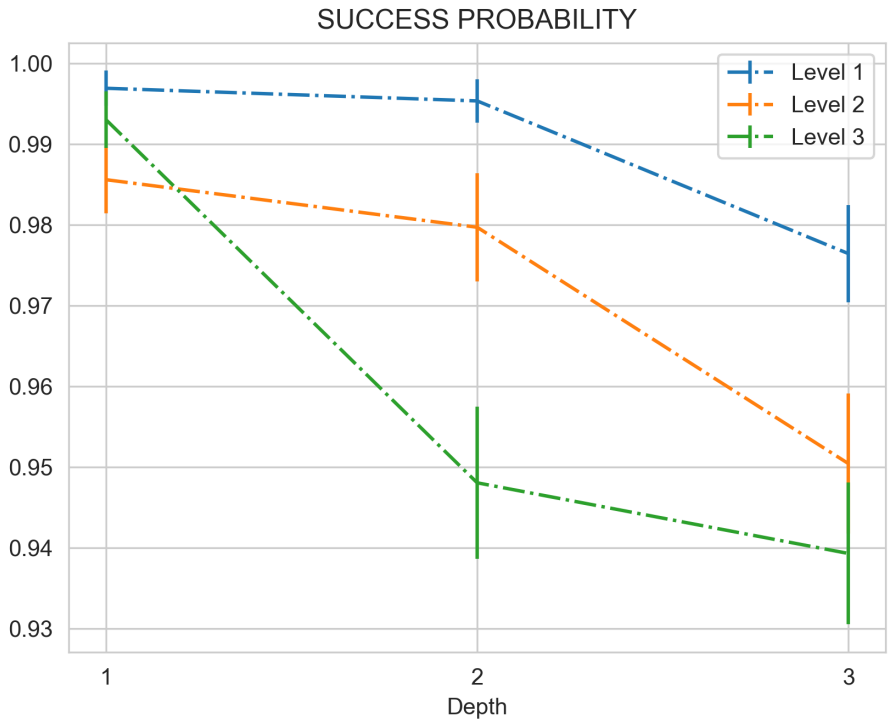

Figure 3.1. Success probability as a function of the 3 levels of the experiment and of the 3 DepthIDs. The figure permits appreciating that participants' performance decreases as a function of levels, but remains stable across levels when the problem planning depth is 1. See the main text for details. Error bars have been assigned according to the standard error of the mean of a binomial distribution.

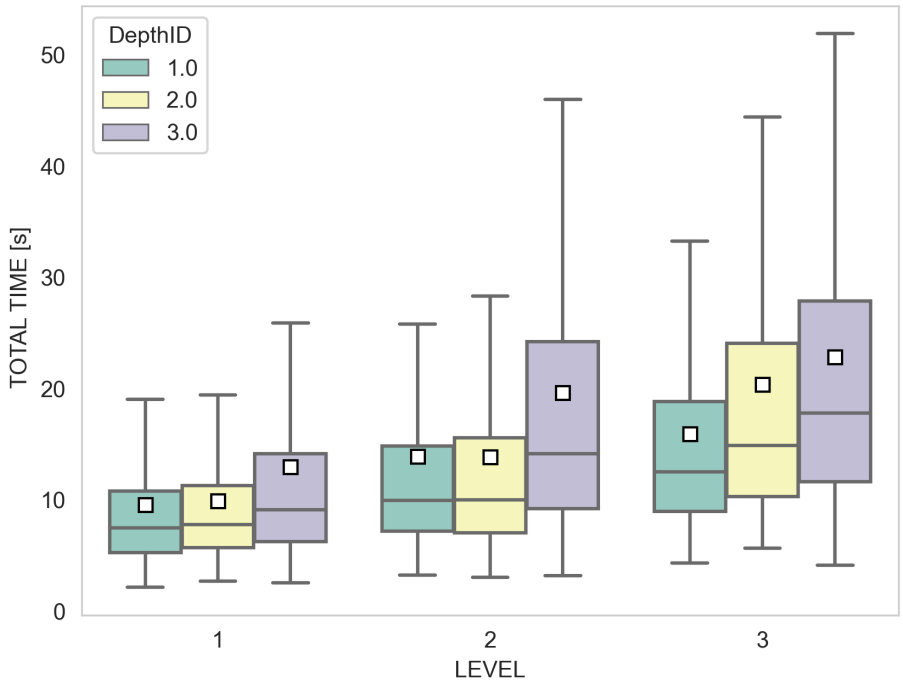

Figure 3.2. Problem completion time (in seconds) as a function of the 3 levels of the experiment and of the 3 planning depths. Results are organized by level and depth. The plot shows the total time, which is bounded to 60 seconds in our experiment

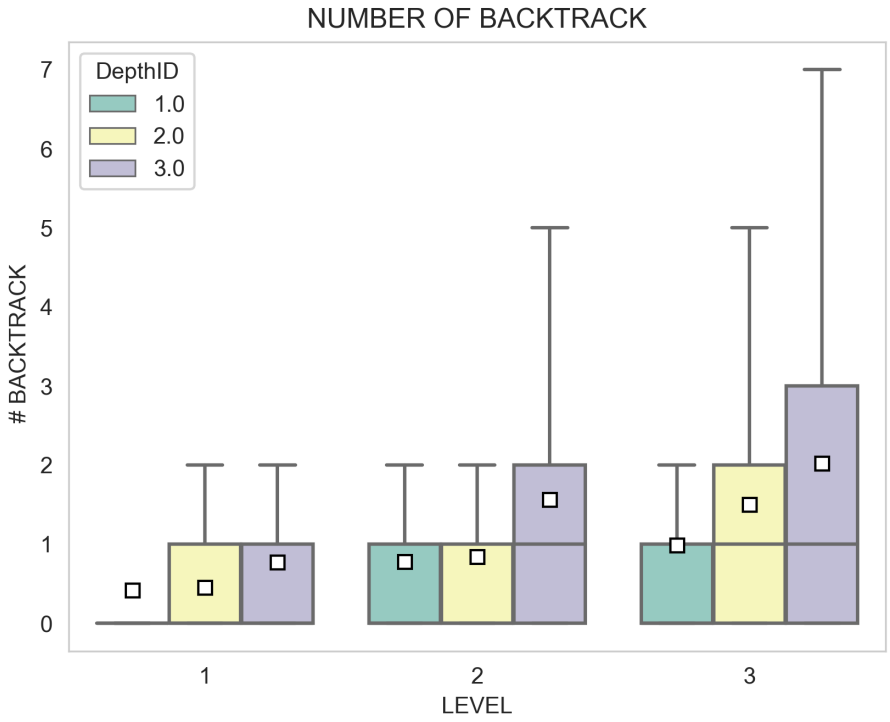

Figure 3.3. Total number of backtracks as a function of the 3 levels of the experiment and of the 3 planning depths. Results are organized by level and depth.

3.3.2 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 3.4. 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.

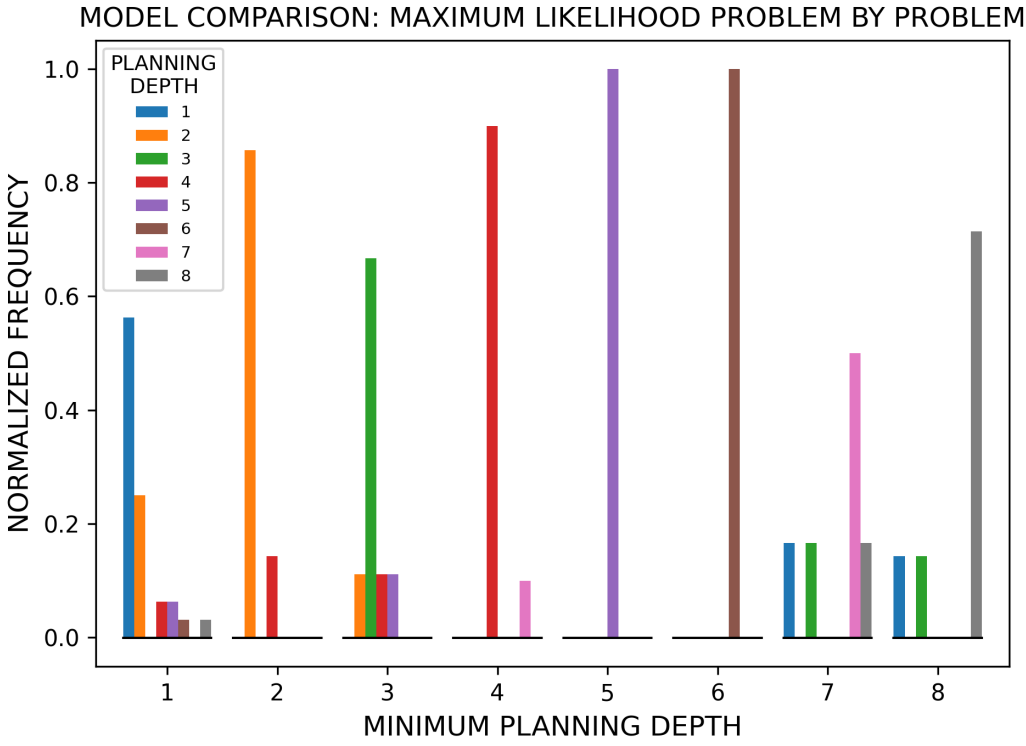

Figure 3.4. Comparison between participants and planning models. The figure shows the number of times that each planner from depth 1 to 8 had the maximum likelihood of the gems collected before the first backtrack by the participants, in the problems of each of the 8 problem groups. Problems are grouped 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 3.5

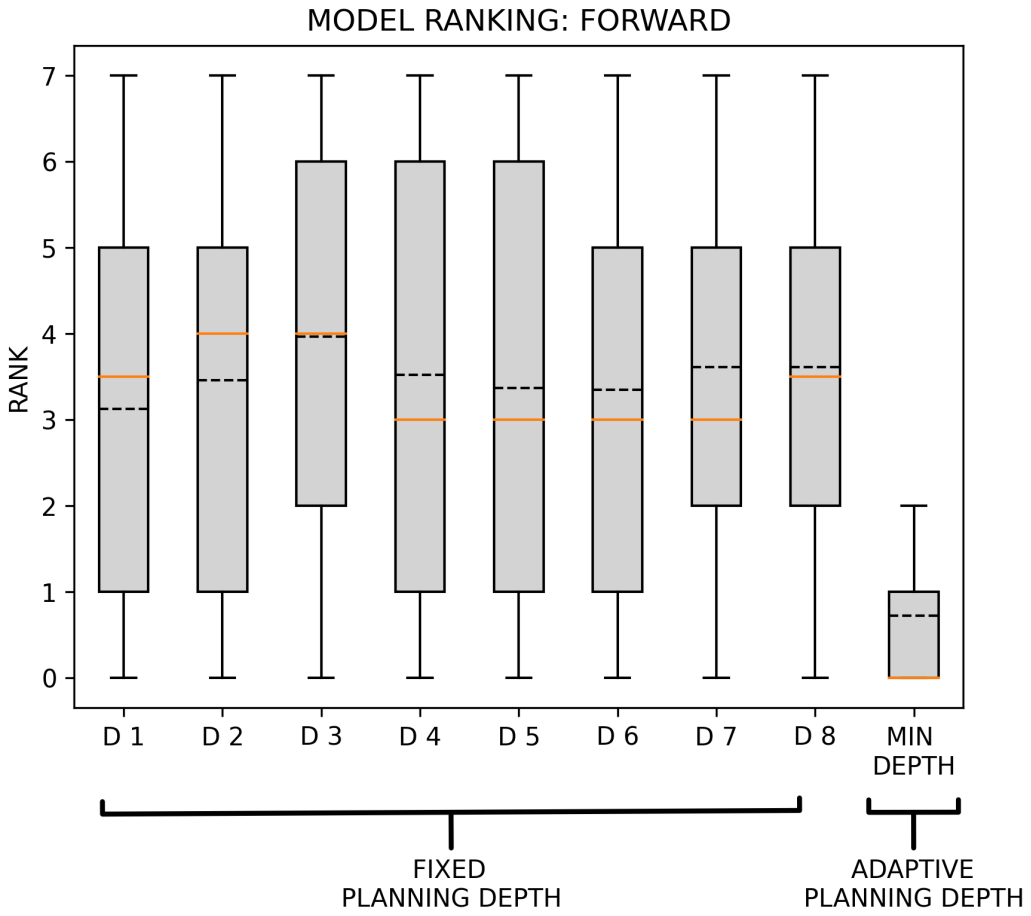

Figure 3.5. Rank values of fixed and adaptive depth planners across the 90 problems. The figure shows the distributions of the ranks of planners using a fixed planning depth (D1-D8) across all problems and of the planner using a planning depth adapted to the minimum required depth to solve a problem (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 is significantly lower than each of the fixed planning depth models 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.

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 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. 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 Figure A.1 and Figure A.2).

3.4 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, et al., 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 3.4. 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 planner that uses a fixed planning depth across all problems (Figure 3.5). This result holds also when considering an additional model that also has the information about the identity of the last gem to be collected, when this information is obvious to participants (Figure A.1).

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 (Balaguer et al., 2016; Q. J. Huys et al., 2012) 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 Traveler 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.

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).

A limitation of this work is 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).

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 app for the task 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.

Chapter 4

Coarticulation in eye-hand coordination as a signature of planning

In this second work we searched for behavioural signatures of planning process in the ThinkAhead task. The experiment was carried out with a computer via mouse controls, while gaze position was acquired with an eye-tracker in a head-stabilized configuration. The kinematic of the eye-hand coordination during the solution implementation exhibited a robust sequential structure consisting of a gaze fixation (of a target node) followed by a reaching movement of the cursor (end-effector of the hand) on the screen towards this target. With this alternated dynamics we decomposed the hand movements in a sequence of *motor plans*. Both within and between the motor plans we detected a significant presence of a *coarticulation* effect, i.e. the modification of a single actions composing a plan depending on the next ones. We found this effect both in the kinematic of the mouse control- while navigating the map problem- and in the fixation positions. This result was corroborated by the information that current fixation provides about the direction of the next fixation. The results suggest that the planning process is not a single-shot process that follows a decision but rather a continuous process where the decision and the action are intertwined and influence each other.

4.1 Introduction

Planning is a covert process and because of this many modelling studies look for similarities between the solutions adopted by humans and the ones adopted by

a model, or, in other words by comparing *what* has been chosen. Response time analysis yield deeper insight beyond the final choice, providing indication of the cognitive effort required. In a popular traditional model family -known as drift diffusion models- it is assumed a sequential process where evidence is gathered before making a decision (Fudenberg et al., 2019; Myers et al., 2022; Ratcliff, 1978; Stone, 1960).

However, in real life, decisions often need to be made simultaneously with actions: as when driving, or playing soccer or cooking. In these cases, the planning process is not a single-shot process but rather a continuous process, where action and decision either act on parallel or form a closed loop, affecting each other (Maselli et al., 2023).

Kinematic analysis of movements offers even deeper insights into the decision process that are not apparent from final choices alone. For example, in a study involving a binary discrimination task using mouse selection, the trajectory curvature has been linked to the stimulus ambiguity, linking the decision to the process unfolding over time. This information would have been unavailable from simple button-press tasks (Barca & Pezzulo, 2012).

This interplay between action execution and planning is not limited to mouse movements (Barca & Pezzulo, 2012; Freeman et al., 2011; Song & Nakayama, 2009). Studies on throwing tasks analyzed the predictability and style of throws by examining whole-body kinematics, revealing a phenomenology richer than the simple final choice (Maselli et al., 2017, 2019, 2022). Overall, kinematic analysis significantly enhances our understanding of decision processes and should be considered in experimental design.

Another possibility is to study the gaze behaviour during the planning process (Hayhoe & Ballard, 2005; Rothkopf et al., 2016). Gaze behaviour can be related to both visual characteristic of the stimulus (Itti & Koch, 2001; Mannan et al., 1997; D. Parkhurst et al., 2002; D. J. Parkhurst & Niebur, 2003), and the planning process itself. Sometimes this can be done by directly looking at what is being watched and with which temporal dynamics (Baldauf & Deubel, 2008; Binder et al., 2023; Najemnik & Geisler, 2005; Pelz & Canosa, 2001; Rayner, 2009; Zhu et al., 2022) in order to determine the underlying strategy and where the attention allocation. This is particularly true during preparatory phases (Crowe et al., 2000; Fiedler & Glöckner, 2012), which are often found in the planning process, and where the gaze is used to scan the environment and to gather information about the problem. Studies have shown that in complex tasks, most eye fixations are on objects immediately needed, but some are on objects needed in the future (Pelz &

Canosa, 2001). Additionally, research on a variation of tic-tac-toe demonstrated that participants' gaze could predict upcoming moves, with accuracy increasing with their expertise In (Huang et al., 2023). Further analysis of gaze behavior in route planning tasks reveals consistent backward and forward planning (Zhu et al., 2022).

Research indicates that saccades, like fixations, are linked to future action planning. In a two-saccade switching task, planning during scanpaths was influenced by changes in visual stimuli. Studies on attention allocation during saccade sequences reveal that saccades are planned simultaneously rather than sequentially, and their amplitude and direction are affected by subsequent saccades in a sequence (Ames et al., 2019; Azadi et al., 2021; Baldauf & Deubel, 2008; De Vries et al., 2014; McPeek et al., 2000; Robinson, 1963).
Thus, if not explicit in the eyes behaviours themselves, the planning process can be inferred by the presence of these modifications induced by the planning process, defined as *coarticulation*. Coarticulation is a well-known phenomenon in the domain of phonology where it has been observed that the same group of letters can be pronounced slightly differently depending on what comes next (Daniloff & Hammarberg, 1973; Fowler & Saltzman, 1993; Whalen, 1990). The same has been found in the domain of motor control (Jerde et al., 2003; Shah et al., 2013), yielding a more general definition as a phenomenon in which the execution of subtasks is influenced by the overall task.
Because of the richness in this kind of data, we decided to perform a second data collection of the ThinkAhead task, in person, where both gaze data and mouse movement (hand end-effector) were tracked and recorded. In the rest of the chapter we will use the expression "eye position" and "hand position" to indicate the projection of the gaze (averaged between the two eyes) and the mouse position on the screen, and by eye-hand distance the distance between these two projections.

By looking at the eye-hand distance over time it seems that the implementation of a path in the ThinkAhead task possess a sequential structure, with a characteristic coordinated dynamics between eye and hand movements shown in Figure 4.1 and Figure 4.2, where the implementation of a solution is realized as an alternation of fixation-reaching movement. A similar situation in a different task is reported in (Zhao & Marquez, 2013).

Obviously this kind of ordered dynamics does not explain all the variability in the implementation of a plan: we excluded in this analysis all the cases where the participant was pausing, or backtracking, because this kind of events are characterized by a dynamics most likely influenced by the uncertainty of the plan which often realizes in an increased variability. However it is also a good starting point to understand the structure of the plan itself.

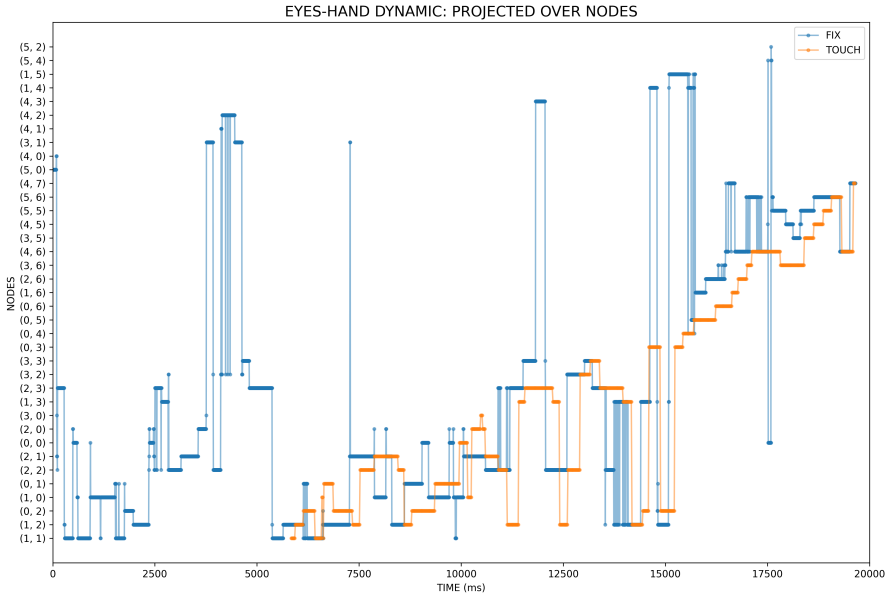

Figure 4.1. A visualization of the eye-hand coordination in a specific trial of the ThinkAhead task. On the y-axis nodes of the graph named after their coordinates, on the x-axis time. Gaze and hand position are collapsed on the closest node of the graph. The blue line represents the gaze position, the orange line the hand position. It can be seen that the eye position is ahead of the hand position most of the times. In the initial phase the eyes scan the graph and the hand is still.

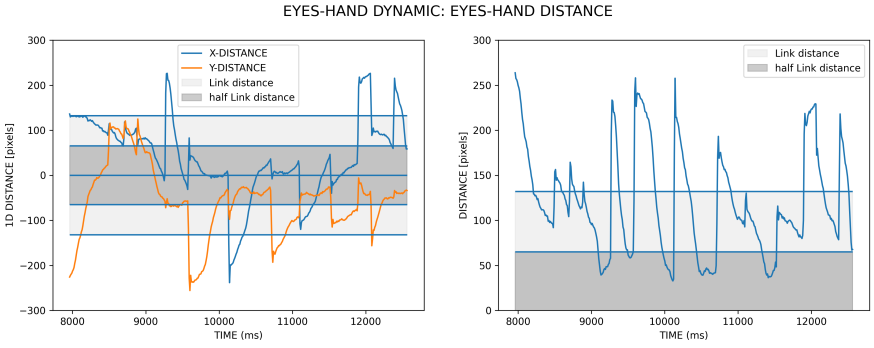

Figure 4.2. The distance between the gaze and the hand position in a trial of the ThinkAhead task (different from the one in figure Figure 4.1). On the left: the difference in coordinates is computed separately for the two dimensions x (blue) and y (orange); shaded areas correspond to the full link distance (lighter) and the half link distance (darker). On the right: the euclidean distance between the gaze position and the hand position.

This alternated dynamics provided a (more natural) temporal segmentation of the hand movements as what happens during a fixation. Thus, looking for coarticulation in this alternated dynamics means looking for: (1) a modification of this movement depending on the next one; (2) the modification for the eye position depending on the next action. In other words we checked if both gaze behaviour and motor control during the plan implementation carried information about future actions.

To preview our results: in both eye movements and hand trajectories, though with different intensity, a coarticulation is detected. Interestingly we also show that the gaze angular position can be used to predict the direction of next fixation, which refers to next motor plan, without using any additional information about the problem or the task.

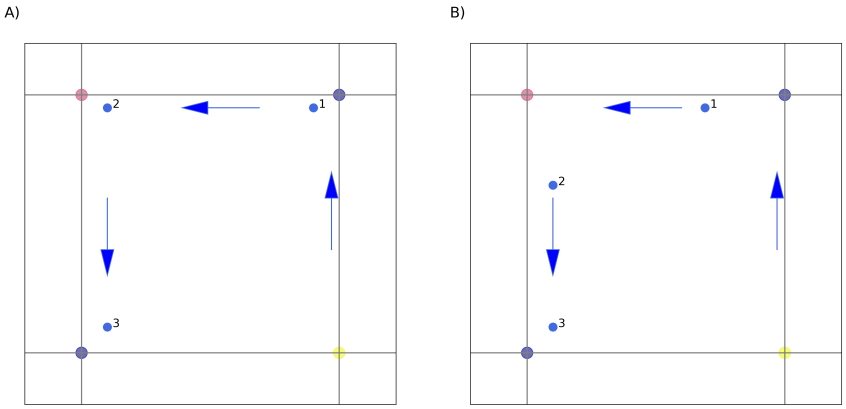

Figure 4.3. An example of the coarticulation effect in the gaze position during a sequence of fixations. Rather than fixating directly the node concluding the motor plan, a slight modification is performed in the direction of the next one.

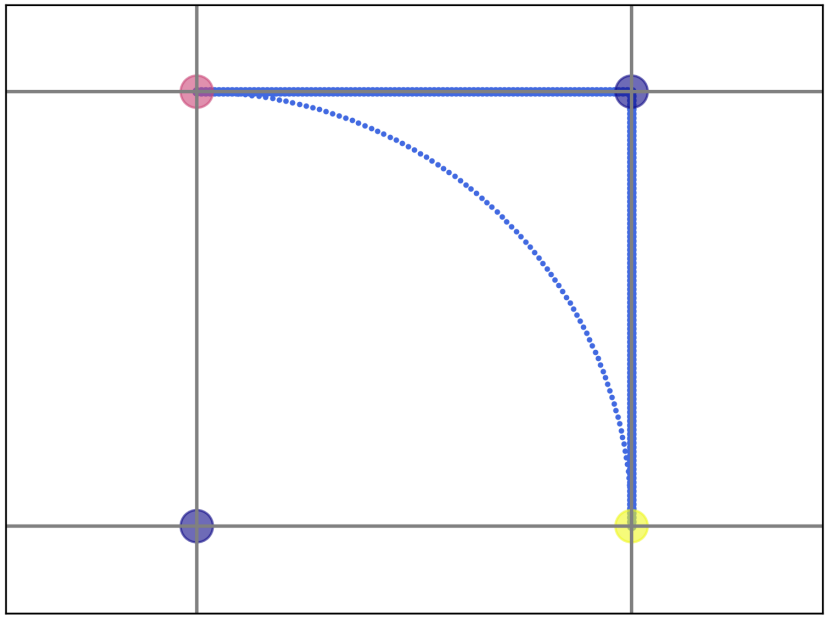

Figure 4.4. An example of the coarticulation effect in the hand trajectory during the execution of a motor plan. Rather than moving directly to the node concluding the motor plan, a curvature is performed in the direction of the next one.

4.2 Methods

Experimental Setup

We used an Eye-Link 1000+ tracker used in a chin-rest desktop mount configuration with a 16 mm lens. The experiment was performed inside an isolated room with no source of light except the IR camera of the eye tracker and the (LED) light of the screen where the task was running. The eye-tracker configuration used was the chin-forehead rest desktop-mount and realized following the indications of the producer: the chin-forehead support was kept at a fixed height such that the eyes of the subjects were approximately at ¾ of the screen height. Subjects seated on a chair with adjustable height. The IR-camera was placed at approximately 50 cm from the eye position of the subject, between the subject and the screen. Subjects could control the movement in the app via a mouse cursor. The synchronization of the ThinkAhead app and the eye trackers was performed via a Python script mainly relying on Pylink libraries (the Python library of the Eye Link) and the PyAutoGUI (a Python library for mouse and visual interactions).

Experimental Procedure

Each subject was first shown the video tutorial (https://drive.google.com/file/d/1RDkY6B_-vM830H_2HgZK1tELSA5N9vJq/view) and given the information about the structure of the experiment: (1) calibration, (2) validation, (3) task divided in three levels of 30 problems each, preceeded by 4 training problems. Subjects were instructed that the total point depended first on the completion of the problem, and then on the velocity of completion. It was also said that the rewards in the task could randomly appear with two colors (blue or red) with the latter case yielding double points. The passage from the trial levels to the first level, as well as the passages from each level to the next, and the end of the whole experiment were indicated by a message on the screen.

For each subject a standard 9+1 points calibration was performed on a grid with the same area of the task problems at the beginning of the experiment. A 9+1 point validation was performed after each calibration. Subjects were asked to move as little as possible once the calibration was validated. The acceptance criterion for the calibration was the one suggested by the producer: a maximum error in calibration of 1° and an average error smaller than 0.5°. Eye tracking sampling frequency was 2000 Hz, while mouse-tracker sampling frequency was 60 Hz.

The duration of the experiment varied significantly with the subject, ranging

from 30 to 60 minutes.

Experimental Subjects

We recruited 31 Subjects (19 M age = 27 ± 4 years, 12 F age=24 ± 5 years) with normal-to-corrected vision took part to the experiment. Participants were free to leave the experiment at any moment. All gave informed consent to our procedures which were approved by the Ethical Committee of the National Research Council All of them were rewarded for taking part to the experiment with a 10€ voucher. All the subjects completed the experiment.

Data Preprocessing: Trajectory resampling

While the sampling rate of mouse tracking was fixed, the movement duration could vary greatly. Thus, the number of points composing a trajectory was not fixed.

In order to standardize the dimensions of the vector representation of trajectories and make them comparable we resampled trajectories with a fixed amount of (2D) points (N = 50). The resampling frequency depended on the velocity profile: so, if in a region of the trajectory geometry original data showed fewer points due to high speed of the movement, the same relative density was used in the resampled trajectory.

We used a linear spline with no smoothness effect, in order to be sure that no additional information about curvatures was added. Notice, however, that in many of the case the choice of the resampling points for the trajectory was an oversampling of the original data (i.e. the resampled trajectory has more data points than the original one). This solution was preferred to a downsampling because it allowed to keep the same amount of information in the trajectory, and to avoid the loss of information due to the downsampling.

It is important to note that the i-th point in different trajectories corresponds to different absolute times but is equivalent in terms of the percentage of trajectory completion. Notice also that in order to have a larger statistics of trajectories to compare, we decided to roto-traslate the trajectories so that the starting point was always the same (the origin (0,0)) and the direction of the first link to be the same (the positive y-axis, or *North* direction).

Spatio-Temporal Principal Component Analysis (st-PCA)

The spatio-temporal principal component analysis (st-PCA) is a method that allows to decompose the variability of a set of trajectories in a set of principal components. It is a generalization of the standard PCA capable of handling the temporal dimension of the data without resorting to a time-series analysis.

The idea behind this method is to treat each point of the resampled trajectory as a distinct variable. The 2D coordinates are flattened in a single vector, resulting in each trajectory of N points being represented as a 2N-vector $(x_1, \ldots, x_n, y_1, \ldots, y_n)$. Stacking all the considered trajectories creates an $N_{trajectories} \times 2N$ matrix, allowing us to perform a standard principal component analysis.

The PCA generates a basis of trajectories, and their linear combinations can accurately reconstruct all the original trajectories with absolute precision. Selecting a subset of this basis enables us to reconstruct the original trajectory with reduced accuracy, as it involves discarding a certain amount of variability present in the original trajectory. This approach allows for the description of the complex trajectory space with a standard method, the PCA, incorporating in a simple way the complex information of the kinematics.

Linear Discriminant Analysis

This low-dimensional representation of the trajectories can then be used to train a classifier to predict the next action.

A widely employed method to do so is the Linear Discriminant Analysis (LDA). The primary objective of LDA is to identify linear combinations of features that optimally differentiate between two or more classes, concurrently minimizing the variance within each class. Its aim is to determine a data projection onto an $n_{classes} - 1$ dimensional space such that the distances between class means are maximized, while the spread (variance) within each class is minimized.

In the feature space, the boundary delineating classes is a hyperplane, representing a linear combination of the features. This method has been successfully applied in the field of motor control, demonstrating efficacy not only in addressing multi-class problems but also in providing a straightforward interpretation of the factors contributing to class distinguishability.

Repeated Stratified K-Fold Cross-Validation

To evalutate the performance of the classifiers we used a standard cross-validation procedure.

Repeated Stratified K-Fold Cross-Validation is a comprehensive technique employed to evaluate models' accuracy. This method integrates two phases. Firstly, in the Stratified K-Fold Cross-Validation phase, the dataset is divided into K folds, ensuring a consistent distribution of class labels across each fold to address imbalances. The model is then trained on K-1 folds and tested on the remaining fold, with this process repeated K times, using a different fold as the test set in each iteration. Subsequently, the Repeated Cross-Validation aspect involves repeating the entire Stratified K-Fold Cross-Validation process multiple times. Each repetition introduces a fresh randomization of the dataset, leading to distinct splits into folds. The model's performance is assessed across all repetitions, providing a robust estimate of its average performance, and accommodating potential variations in the dataset splitting process. This approach enhances the reliability of evaluating a model's generalization capabilities.

Data Preprocessing: distinguishing a general and a restricted case of eye-hand coordination

As anticipated in the introduction, from a preliminary analysis of the eye-hand distance we noticed an alternated dynamics of eye-hand coordination (sequences of fixations and reaching movements). However, since the aim of the study was looking for the information that this dynamics carried about future and in many occasions the eye performed a saccade towards the next target before the current reaching movement was completed, we decided to distinguish two cases: a *general* and a *restricted case*.

In the restricted case we only considered cases where the fixated node was the last reached node during the fixation (in the sense that the next saccade was performed only once the hand reached the gaze position).

Data Preprocessing: splitting data in within and between motor plan condition

Once assumed the segmentation of the hand dynamics into motor plans, we looked for coarticulation effects in two different conditions. The first case is *within* a motor plan: this case requires us to consider all the fixations during which at least two

links have been crossed. The trajectory to select the first node trains the classifier and the second one will act as test to be classified. Notice that while for the training set the full trajectory is used, for the test set it is discretized into four possible directions (corresponding to the cardinal directions of the lattice).

Additionally, fixations performed during a pause or a backtrack were excluded. The second case is *between* two motor plans: during the first fixation a single link had to be crossed, while during the second fixations at least one link. These coupled fixations had to be performed during a movement where no pauses or backtracks occurred.

4.3 Results

The four main results of this second experiment are: (1) the identification of a regime of eye-hand alternated dynamics in the implementation of a plan; (2) the estimation of the coarticulation effect in the kinematics of the mouse control; (3) the estimation of the coarticulation effect in the gaze position; (4) the predictive relation between current fixation and next fixation.

4.3.1 Eye-Hand Alternated Dynamic

If we restrict ourselves to the case of fluid movements (i.e. movements that are not interrupted by pauses), the dynamics of eye-hand coordination exhibits an alternating structure, composed of sequences of fixation-reaching movements. As shown in Figure 4.1, the eye position is "ahead" of the hand position most of the times, and the gaze position is mainly on the target node, and is then followed by a reaching movement of the hand towards the gaze position.

To test the statistical significance of a regime in which new fixations are performed towards a further position, and then the hand reduces its distance from the gaze position during the fixation, we examined two quantities: (1) the difference between the eye-hand distance at the end of a fixation ($d_t(eye, hand)$) -i.e. slightly before a new saccade is performed- and the eye-hand distance at the beginning of a new fixation ($d_{t+1}(eye, hand)$), and (2) the derivative of this eye-hand distance over time after the new saccade is performed.
The difference in eye-hand distance computed between consecutive fixations is defined as:

$$\Delta d_t = d_{t+1}(eye, hand) - d_t(eye, hand) \tag{4.1}$$

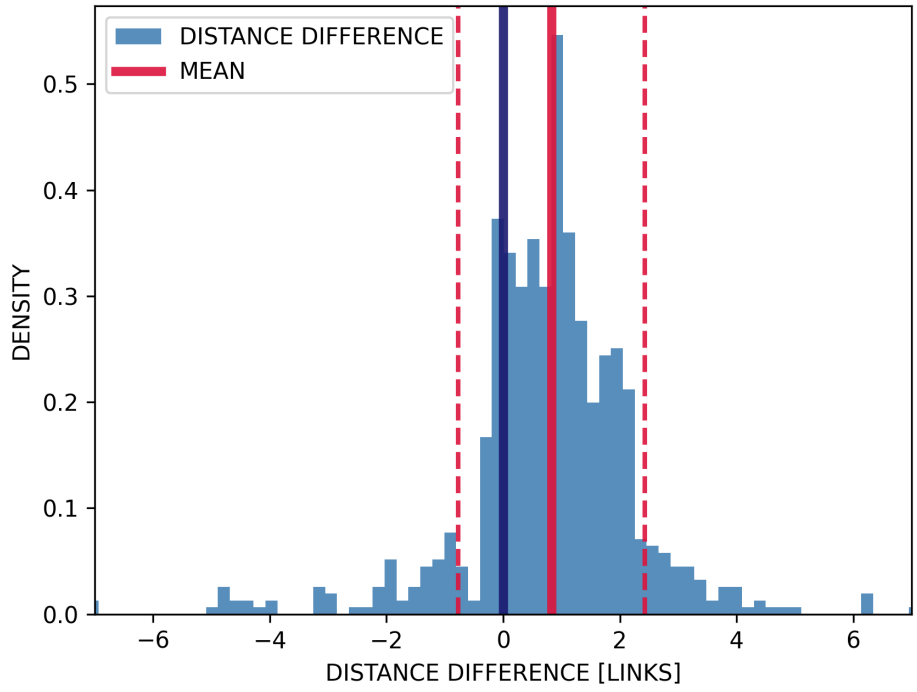

Figure 4.5. The difference in the eye-hand distance between two consecutive fixations. The distance is computed as the euclidean distance between the two positions. Positive values of this distance difference mean that the gaze moved further from the current hand position with the next fixation, while negative values mean that the gaze shortened the distance from the current hand position with the next fixation. The (solid) red line is the mean of the distribution; the (dashed) red lines identify an interval of two standard deviation; the (solid) blue line is set at zero. The mean of the distribution is significantly positive, meaning that the next fixation is usually done towards a farther position with respect to the current position.

Positive values of this distance difference mean that the gaze moved further from the current hand position with the next fixation, while negative values mean that the gaze shortened the distance from the current hand position with the next fixation. As shown in Figure 4.5, the difference in eye-hand distance computed between consecutive fixations is on average significantly greater than zero ($p_{value} < 10^{-3}$). Thus, when a new fixation is performed, a node which is further with respect to the current position is usually selected.

Then, we tested if the eye-hand distance Δd decreases during the fixation, as expected from a reaching movement. To do that we computed the temporal derivative of Δd during the fixation duration: a negative value in the derivative is associated with a reduction over time of the eye-hand distance, while a positive value is associated with an increase over time of the eye-hand distance. Note that in order to

compare distances during fixation of different duration, we first defined a normalized time $\bar{t}$ given by the absolute time t shifted to start from 0 and normalized by the temporal interval duration:

$$\bar{t} = \frac{t - t_{start}}{t_{end} - t_{start}} \tag{4.2}$$

where t_{start} and t_{end} are the start and end of the fixation, respectively. $\bar{t}$ will indicate the percentage of trajectory completion during current fixation (or, in other words: the percentage of fixation duration), and will be comprised between 0 and 1. Once computed the derivative the eye-hand distance difference with respect to this normalized time $\bar{t}$, we can test the significance of the derivative sign (and thus, the significance of the increasing/decreasing behaviour of the eye-hand distance) at each percentile of fixation completion. As shown in Figure 4.6 the time derivative of this distance is negative throughout the fixation duration, implying that, on average, the hand reduces its distance from the gaze position, as in a reaching movement. The significance of the negative sign of the derivative is confirmed throughout the whole fixation duration ($p_{value}(t) < 10^{-3}, \forall t \in [0, 1]$).

Consequently, in this work we decided to segment the events in the interval of time defined by start and end of fixations and distinguish motor plans happening within a fixation and between two fixations.

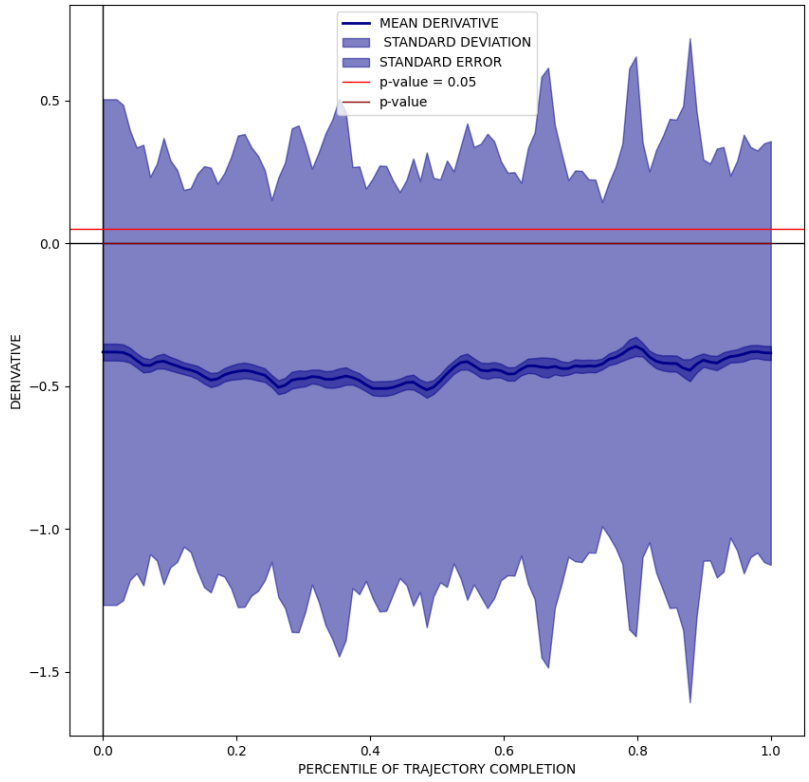

Figure 4.6. The temporal derivative of the distance between the gaze and the hand position during a fixation over a normalized time (percentile of the fixation duration). The blue (solid) line is the mean value of the derivative over time; the (smaller) shaded area is the standard error of the mean, and the (wider) shaded area is the standard deviation from the mean; the (darker) red line is the estimated p_{value} at each time, and the (lighter) red line is a threshold $p_{value}(t) = 0.05$. The fact that the derivative is significantly negative throughout the fixation suggests that the on average after a fixation is performed the hand reduces its distance from it, in a reaching movement.

4.3.2 Spatio-Temporal Principal Component Analysis (st-PCA) reveals meaningful structure in the kinematics

The explained variance of the original data is already good (See Figure A.5 for the cumulative explained variance) when few components are considered, as shown in Figure 4.7. Remaining analysis make use of 10 PCs (however: the main results are tested when varying the number of components to characterize and ensure robustness)

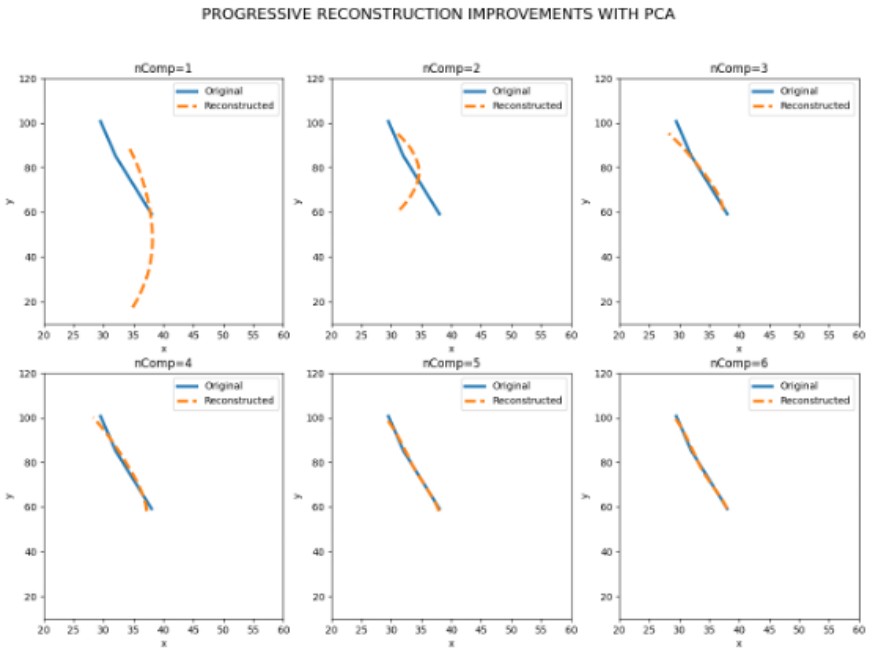

Figure 4.7. Reconstruction of the trajectory from the first 6 PCs in the general case within fixations. The original trajectory is shown in blue, the reconstructed one in orange.

Figure 4.8 shows the projection coefficients of the principal components in the original space (i.e., the 2D 50 points-resampled trajectories, with the first 50 points being the x coordinates and the remaining 50 points being the corresponding 50 y coordinates). Interestingly, not only the higher the number of components, the higher the complexity of the associated movement, but there is also a clear alternation in the relevance of the x and y coordinates in the projection coefficients, which segregate their importance into different components.

Looking at these projection coefficients we see that the first two components mainly describe an offset with respect to the average, with the second component that seems to show a mild dilation/contraction effect along the y-axis. Third and fourth component represent a curvature in the trajectory where there is a crossing of

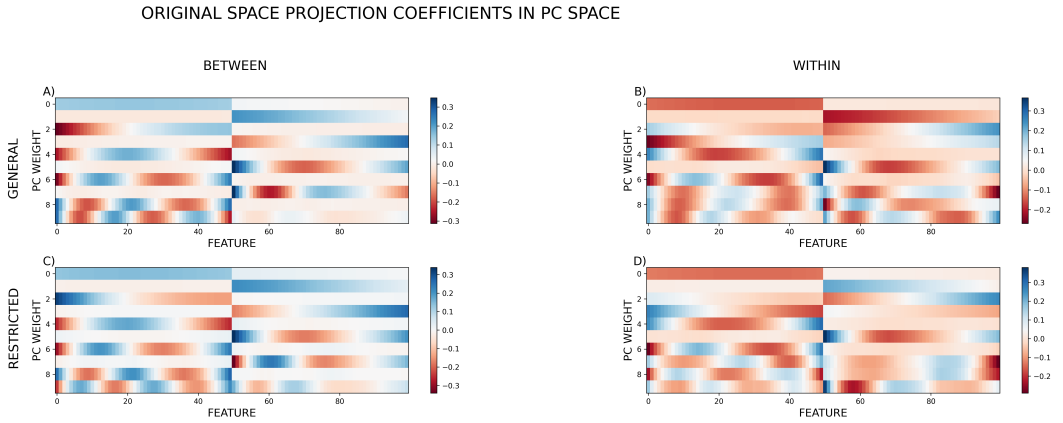

Figure 4.8. PCA coefficients for the first 10 PCs in the general case within fixations. The values are the coefficients of the linear combination of the resampled trajectory that allows to reconstruct the original trajectory.

the vertical axis. Fifth and sixth component introduce curves where the beginning and the ending of the trajectory bend towards a direction, but the middle part of the trajectory moves toward the opposite. We could say than that in these first components there are two main movement strategies: one -third and fourth components- that looks for the shortest path (and thus bend the trajectory towards the next target), and one -fifth and sixth components- that looks for a smoother path, despite the biggest length. Higher components show a higher complexity, with more S-shaped trajectories.

Interestingly, in the within-fixation case - and in particular in the seventh and eighth components - we see that the segregation effect is milder or lost, and replaced by a PC that mixes the x and y dimensions. In the case of these last components however, the corresponding PCs in the between fixations case can be recovered via a simple sum-difference operation. Despite these differences, Figure 4.8 shows that the same structure of the principal components is shared among the cases considered (Between/Within X Restricted/General). Thus, we decided to perform the same analysis on all the data, and use the resulting principal components as compressed base for the trajectory description in the following classification task. The projection of these components into the original space is shown in Figure 4.9.

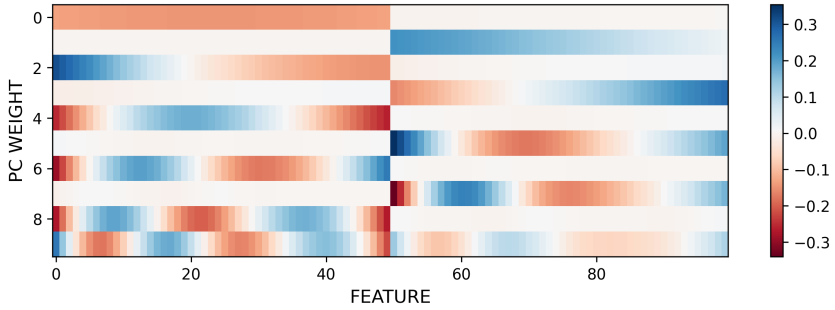

Figure 4.9. Projection of the first 10 components in the original space.

4.3.3 Detecting Coarticulation in the Hand Kinematics and Eye Position

Linear Discriminant Analysis

In Figure 4.10 there is an example of LDA coefficients, where we have projected the PCs over the LDA space. The values show how much each principal component allows to discriminate the three labels considered in the classification task.

The coefficients shown here are for four LDA models trained over all the data of that case. This model is obviously prone to overfitting and is not used for the cross-validation, so the corresponding values should be considered simply as an example. Let's note that the discriminative power increases with the number of components considered, while the unexplained variance decreases with them. This means that the overall coarticulation effect will result from the balance of these two factors.

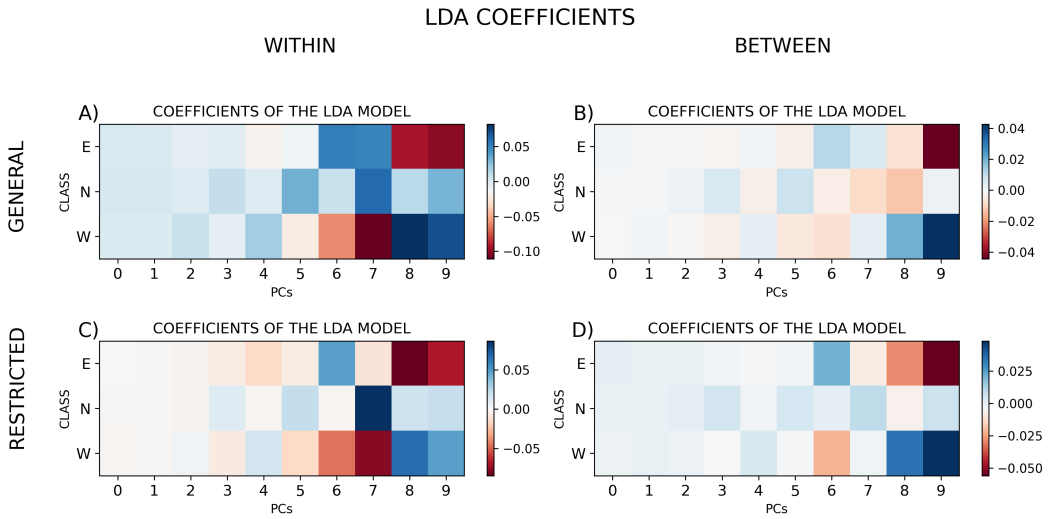

Figure 4.10. LDA coefficients for the first 10 PCs in the general case within fixations. The values are the coefficients of the linear combination of the PCs that allows to discriminate the three classes.

The Classification Validation: Accuracy Estimation, K-fold Cross Validation and Confusion Matrices

In order to assess the statistical significance of the LDA model accuracy in the different cases, we used a repeated stratified K-fold cross-Validation procedure (number of folds = 3, number of replicas = 10). In Figure 4.11. We computed the average classification accuracy for LDA models on the different cases. We used the LDA for both the eye position with respect to the origin of motif, and the trajectory projected on the compressed PC space (10 components). In all the cases we balanced the classes' volume by a downsampling procedure, so the random case we had to compare models' accuracies was $\frac{1}{n_{classes}}$, with $n_{classes}$ = 3 and thus 0.33. In Figure 4.11 we see that: (1) accuracy is always higher than the random model; (2) accuracy is higher within a fixation, rather than between different fixations, for both the cases (general/restricted) and both the predictive variables (eye fixations/ hand trajectory); (3) the eye position provides more predictive power than the kinematics of the trajectory.

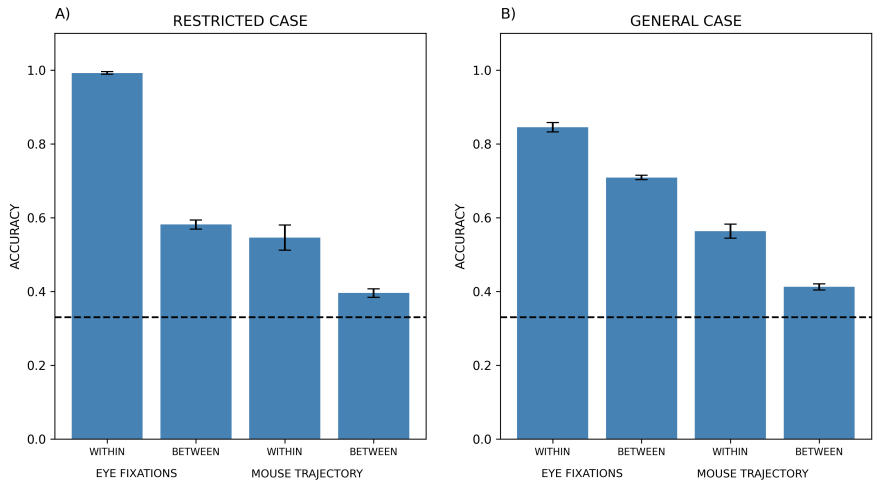

Figure 4.11. Accuracy of the LDA models in the different cases. The error bars represent the standard deviation of the accuracy across the 10 replicas of the cross-validation.

We also tested the robustness of the coarticulation effect when considering a different number of principal components in the trajectory description, varying from 1 to 10. Figure **??** demonstrates that coarticulation effects in the *within* cases become significant (i.e., exceed chance level) when considering more than the first two components (the 'offsets'), whereas in the *between* cases, even the offsets contribute to discrimination.

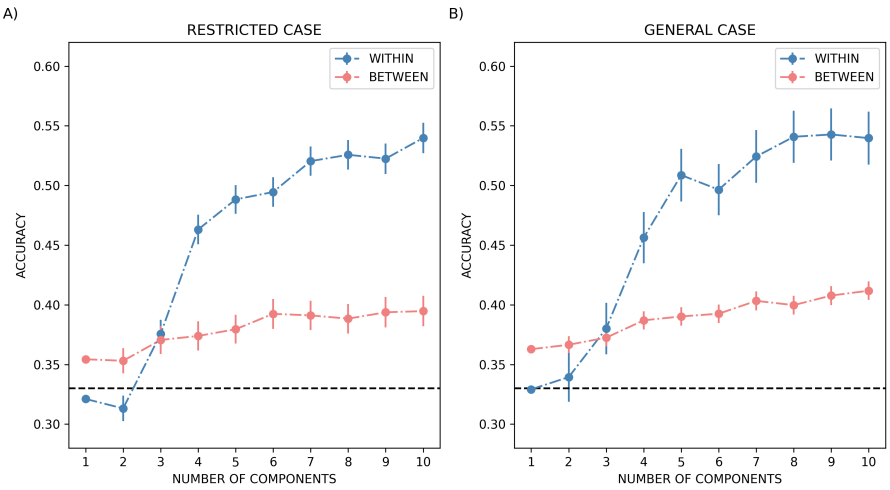

Figure 4.12. The accuracy for LDA classifiers trained on trajectories compressed to a varying number of principal components (from 1 to 10) for (A) the restricted case and (B) the general case.

Confusion matrices give us more information into how the quality of the classification. Between fixations, the hand trajectory classification is not able to classify the north direction against the west and east direction, and it turns out to be just a random classifier. In all the other cases the diagonal shows higher values than the off- diagonal terms.

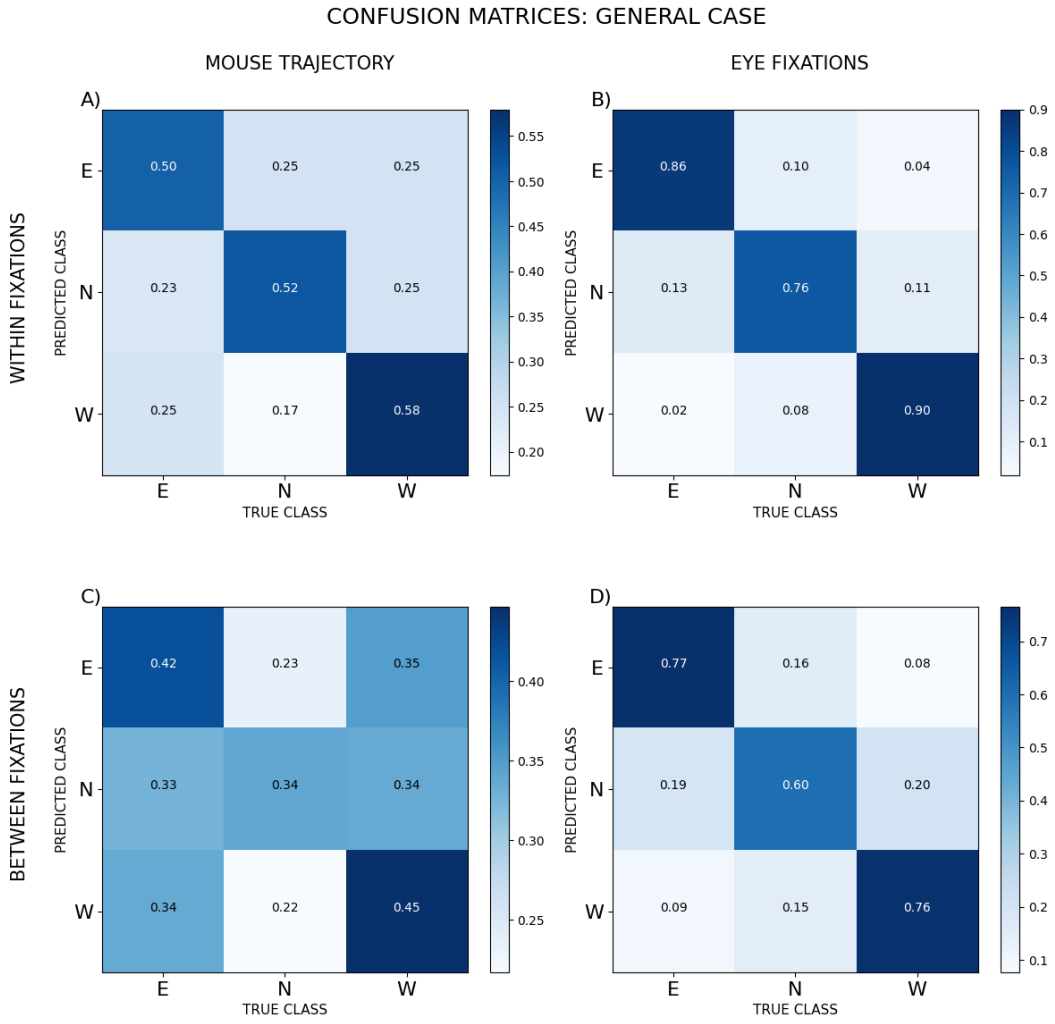

Figure 4.13. Confusion matrices for the LDA models in the general case. The values are the average of the confusion matrices across the 10 replicas of the cross-validation.

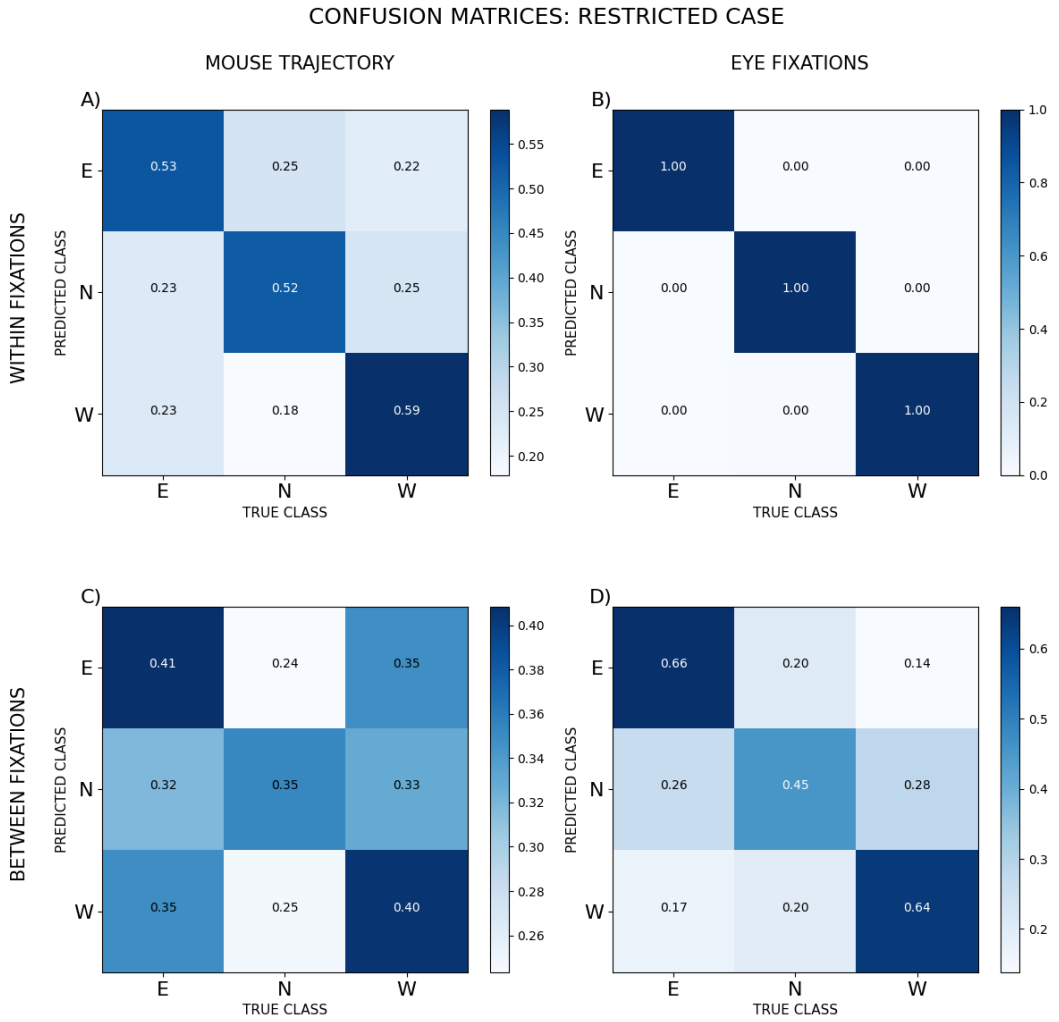

Figure 4.14. Confusion matrices for the LDA models in the restricted case. The values are the average of the confusion matrices across the 10 replicas of the cross-validation.

4.3.4 Predicting next fixation: coarticulation in the eye angular position

We analyzed more in depth the information shared between the current fixation and the next one. As previously said: if the motor implementation phase was just a fixation-reaching sequence, the gaze position could simply fixate very closely (apart from physiological noise), the ending node of a motor plan. However, in many cases the eye position tilts slightly towards other direction. Because of this we are able to define an angle that span from a reference point -for example the horizontal axis- and the direction connecting the eye position and the ending node of a motor plan.
We thus fitted the relation between the current fixation angle and the next fixation angle.

Figure 4.15 shows the relation in a Cartesian plane. However, since both of the variables are angular, the map from one to the other is defined over a torus. To fit the relation between these two angles we used a circular-to-circular regression (as formulated in (Jammalamadaka & SenGupta, 2001)). The regression fits sine and cosine of the dependent angle with a trigonometric polynomial of order m of the independent angle. This regression comes with a natural statistical test to assess the significance of increasing the order of the polynomial.

To see if the angular position of the current fixation carried information about the next one, we compared the accuracy of the model used to regress the current fixation angle to the next fixation angle, and to two random models. The first one used a random point around the nodes that had not been visited; the second one used a random point in an interval around only the rewards that had not yet been collected. The importance of these two random models is that they do not rely on planning information (such as for example future path), but rather on perceptual information constantly available to the subject (position of available nodes, position of missing rewards).

To test the predictive power of these three models we used a classification task: we discretized the angles into 8 bins of equal angular span. In this estimation accuracy it was not used a downsampling procedure because classes were (approximately) balanced: the null case was 0.126 versus 0.125 of a perfectly balanced case $\frac{1}{8}$.

To estimate the accuracy of the model regressing the current fixation angle to the next fixation angle we used a 3-fold cross-validation procedure repeated 50 times. At each folding we performed a new circular to circular regression -thus estimating first the necessary order of the trigonometric polynomial- discretized the predicted output, and thus we computed the accuracy of the model.

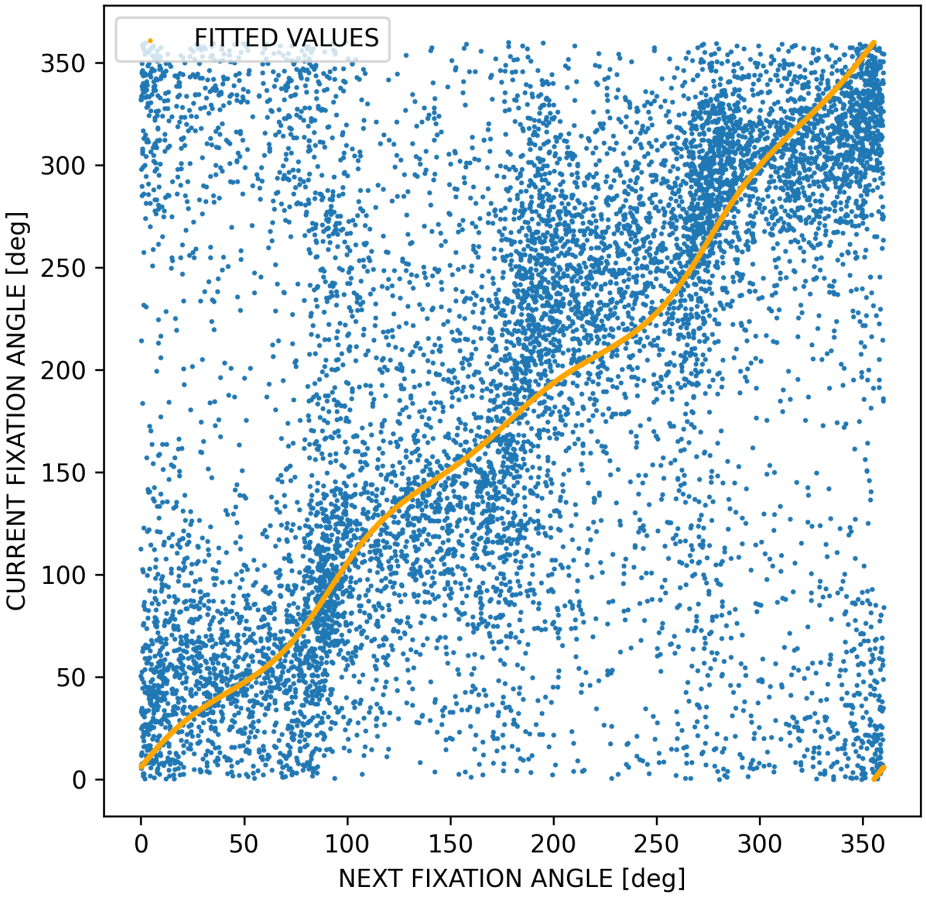

Figure 4.15. Relation between the current fixation angle and the next fixation angle in the general case. The orange line represents the fitted values of the circular-to-circular regression.

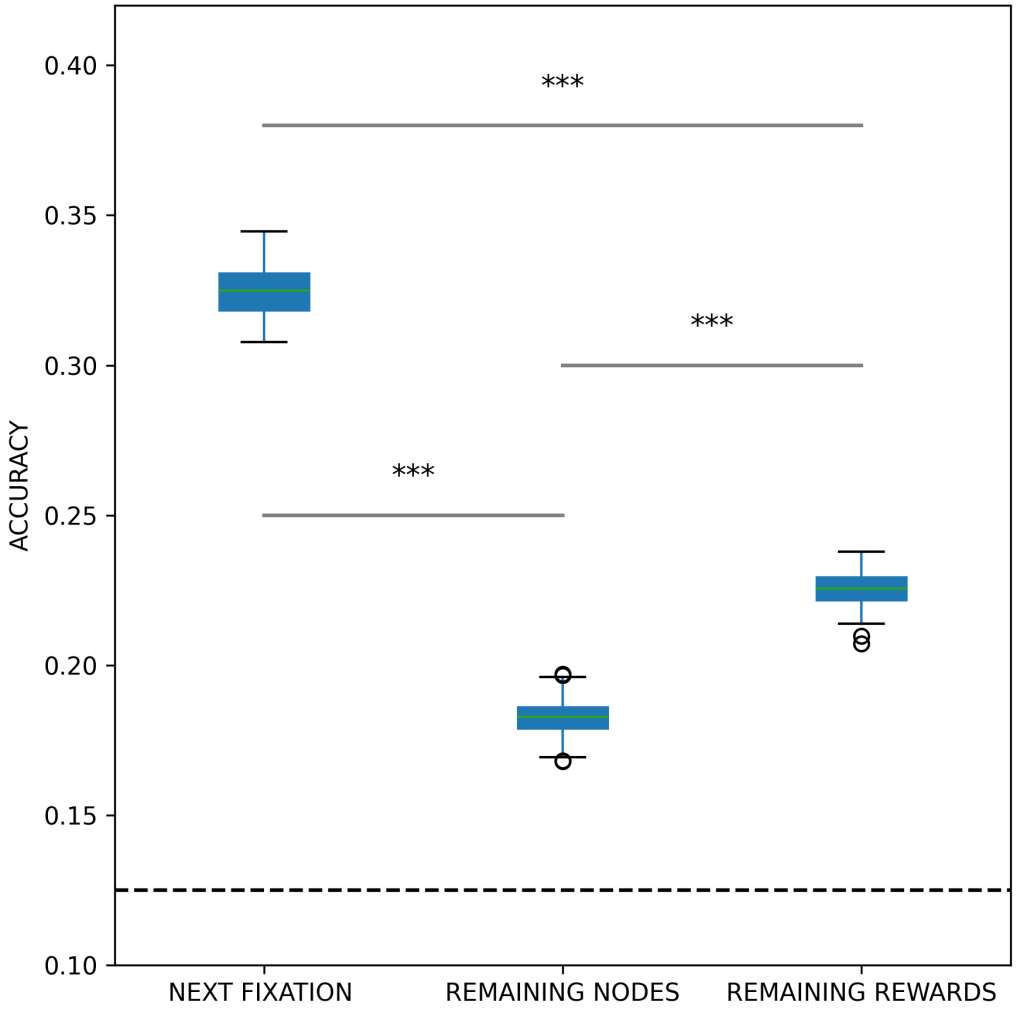

Figure 4.16. Accuracy of the circular-to-circular regression model in the general case. The boxplots represent the accuracy of the models. The dashed line represents the accuracy of the uniform chance level.

For the random models we additionally considered 30 sets of randomly sampled variables (unvisited node for the first model, uncollected reward for the second one), and for each of them we performed a 3-fold cross-validation procedure repeated 10 times. The accuracy of the model was then computed as the average of the accuracy of the 30 instances of the models.

As shown in Figure 4.16, the circular-to-circular regression model is able to predict the next fixation angle with an accuracy significantly higher than the random models ($A_{NextFixation} = 0.350 \pm 0.008, p_{value} < 10^{-3} A_{RemainingNode} = 0.183 \pm 0.006, p_{value} < 10^{-3}, A_{RemainingReward} = 0.225 \pm 0.006, p_{value} < 10^{-3}$). This result is consistent with the idea that the gaze position is influenced by the future actions,

and thus that the implementation of a plan is not a single-shot process but rather a continuous process that affects the implementation of the plan itself. Note also that all the three models have an accuracy higher than the uniform chance level.

4.4 Discussion

Planning is a covert process, and thus it is difficult to study it directly. However, as long as the output of a planned decision takes place in the world, we can study both the decision taken and how it has been performed, in the attempt to track the planning process. In this second experiment we expanded the paradigm by considering not only the solution to the task chosen by participants, but also the eye-hand dynamics in doing it.

A preliminary analysis of the dynamics of their coordination showed the presence of an alternation between a fixation and a reaching movement, which was found to be consistent during fluid movements (i.e., movements that did not present pauses).

This structure was thus employed to segment the data in the interval of time defined by the start and the end of a fixation, and to distinguish motor plans happening within a fixation and between two fixations.
The interesting variables that we studied in this segmentation were the kinematics of the hand trajectory and the gaze position. The kinematics of the hand trajectory was analyzed via a spatio-temporal PCA, which revealed a meaningful structure in the kinematics of the movements, and two main possible movement strategies, compatible with the idea of a trade-off between the shortest path and the smoothest path.

Also, when using this comprehensive low-dimensional description of the trajectory, we found that the hand trajectory was able to predict the direction of the next fixation with an accuracy significantly higher than the random case.
This result is a signature of coarticulation and of the idea that when planning a movement, the implementation of the different parts of a plan is influenced by the future actions.

We also looked for coarticulation effect in the gaze behaviour, and found that it was able to predict the next action with an accuracy significantly higher than the random case.

Finally, we analyzed more in depth the information shared between the current fixation and the next one. We did this by regressing the current fixation angle

to the next fixation angle, and to two other random models based on perceptual information only. We considered a classification task where the angles were discretized into 8 bins of equal angular span and tested the accuracy of these three regression models. We found that the circular-to-circular regression model was able to predict the next fixation angle with an accuracy significantly higher than the random models.

Globally, these results not only confirm the idea that the implementation of a plan is affected by the nature of the plan itself - rather than being conceived as a simple sequence of actions- but it also expands the repertoire of cases where this property of coarticulation has been found.
There are many limitations of the current study that need to be addressed in future works.

The first one is that the paradigm adopted to segment the hand trajectory into motor plans associated with different fixations does not explain all the phenomenology encountered in this task. For example there are cases where the gaze goes back and forth from the current target and future nodes, during the reaching phase. In these cases, rather than a coarticulation there is an explicit prospection of the future actions.

Secondly, and even more important, in this work we have neglected the extremely interesting phases of explicit scanpath that occur during the initial planning phase and during the pauses after that. This would provide valuable insights into the strategy of information foraging that the participants adopt during the planning phase.

Finally, for simplicity of analysis, we had to reject all the cases where the participants trajectories missed the collider of a node when reaching a target. We observed that very often these cases comprised movements where the coarticulation effect was more pronounced and trajectories modifications were exaggerated, so including them in future analysis would provide a stronger result.

Chapter 5

The Determinants of Backtracking: Evidence against continual replanning

This chapter investigates the mechanisms underlying the commitment to plans and the triggers for backtracking during the ThinkAhead task. By analyzing recordings of participant actions and visual feedback, the study reveals how individuals manage their plans in response to task demands and perceptual cues. It was found that backtracking primarily occurs after strong visual feedback indicating the unsolvability of the current plan, suggesting that participants do not continuously monitor their plans but rely on perceptual triggers to revise their strategies. Here, we also propose a taxonomy of possible backtracking signals, distinguishing between cases that requires only visual information and other that require some sort of planning. These findings highlight the dynamic interplay between perception and action in planning and decision-making, contributing to our understanding of how people detect and correct errors.

5.1 Introduction

In previous chapters we described how difficult and resource-consuming planning can be (Anderson, 1990; Russell & Norvig, 1995; Simon, 1957, 1996; Wiener et al., 2009), and what are the main strategies to face these limitations (Alterman, 1988; Balaguer et al., 2016; Gershman et al., 2015; B. Hayes-Roth & Hayes-Roth, 1979b; Q. J. M. Huys et al., 2015, 2012; Keramati et al., 2016; Kuperwajs et al., 2024; Pezzulo et al., 2013; Wilensky, 1981).

In a condition of perfect and exhaustive planning, the agent would not need to replan its actions but, under the (reasonable) assumption of bounded rationality, humans might prefer not to make a full plan from the beginning, but rather make partial plans and then replan halfway (Callaway et al., 2022; Newell, Simon, et al., 1972). This leads to a secondary problem in the planning process: if a full plan cannot be made, how much do I monitor the plan, and how frequently do I replan?

We consider multiple hypotheses about when (and how often) participants replan. One possibility is that in parallel to task execution, participants constantly monitor and revise their plans (continual revision). Another possibility is that they monitor and revise their initial plans intermittently, e.g., at key subgoals (intermittent revision). A last possibility is that they only revise their initial plans when they observe an explicit message of error, such as the fact that they reached a dead-end. In this last case, participants would not fully solve the planning problem in their minds, but rather they would use a strategy of *cognitive off-loading* (Kirsh & Maglio, 1994).

In this section, we show that in this task participants do not monitor their plans frequently, but rather they rely more on perceptual feedback signaling that their initial plan was erroneous (e.g., because the initial plan reached a dead-end). This suggests that the simple unsolvability of the task is not the only trigger for backtracking. Thus, we proposed a taxonomy to help explain what might be the triggers for backtracking and found that in more than half of the cases participants backtrack in conditions of perceptual feedback of errors.

5.2 Methods

5.2.1 Participants and Experimental Setup

For this analysis we analyzed the data from the second experiment. See section 4.2 for the details about the participants and the experimental setup.

5.2.2 Determining the task solvability

In the first analysis we computed the number of steps that separated the moment where the task became unsolvable and the moment where the backtracking started. The task solvability was determined with a model-free approach, by checking if the current path selected by the participant was part of a solution.

5.2.3 Creating a taxonomy of backtracking events

The taxonomy is shown in Figure 5.1 and comprises seven categories, hierarchically organized in three main groups.

The first group contains a single category that includes backtracking events that were not necessary (*solvable*).

The second group includes three categories: unsolvable, double dead-end, prospected proximal bisection. By *unsolvable* we indicated the generic cases that did not fall into any other category; *double dead-end* indicates cases where the current path selected by the participant has made the task unsolvable by creating two rewards which are dead-end (and thus for the task rules are mutually exclusive); the *prospected proximal bisection*, that includes the cases where the path has only one possible development and this lead to the bisection of the graph. This is not properly a planning phenomenon but rather a short term prospection, because only one possible path is to be selected.

The last group includes three categories: *proximal bisection*, *bisection* and *no more rewards*. These categories are characterized by the fact that there was visual feedback of the non-solvability of the task that did not require any form of planning. The *proximal bisection* category includes the cases where in the current position any choice but the backtracking will lead to a clear bisection of the map (while keeping rewards on both the (dis)connected components). The *bisection* category includes the cases where the map has been already bisected and the are rewards in both the (dis)connected components. The *no more rewards* category includes the cases where the participant cannot reach any of the remaining rewards.

Categories are exhaustive but in principle are not mutually exclusive. However, we decided to consider that an event could belong only to the most restrictive category (e.g. a *no more rewards* is obviously also an *unsolvable* case, but we decided to consider it only as a *no more rewards* case).

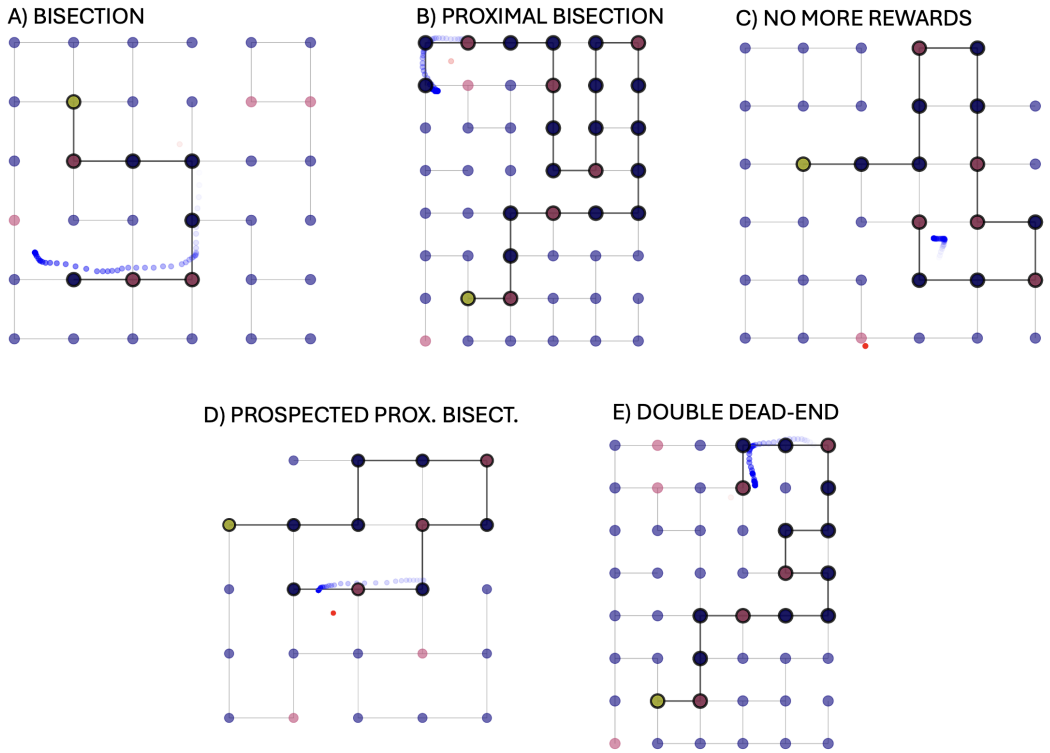

Figure 5.1. Example of the categories of the taxonomy. The figure an instance of the cases making part of the taxonomy. The yellow node is the start location, the red nodes are the rewards. The blue nodes are the nodes that have been visited. Nodes that darkened belong to the path selected by the participant so far. The light blue dots are the trace of the hand trajectory. The figure shows the categories: A) *Bisection*, B) *Proximal Bisection*, C) *No more rewards*, D) *Prospected Proximal Bisection*, E) *Double Dead-end*

5.3 Results

In this work we discussed different hypotheses concerning the monitoring of the plan. First we extrapolated participants' behaviour to understand how long after having made the task unsolvable they started backtracking. Then we proposed a taxonomy to help differentiate the error signals that might have triggered the backtracking event.

5.3.1 Continual monitoring of a plan or visual feedback?

We estimated the distance (intended as number of links selected) from the moment that the task became unsolvable (because of the path selected by the subject) and the moment that the backtracking was started. The unsolvability condition is a model-free property since it is defined by the task rules.

Figure 5.2 shows the distribution of the histogram of this distance. The fact that the distribution slowly decreases after the zero, means that people do not immediately recognizes the unsolvability of the task, but rather keep ongoing building their (wrong) solution.

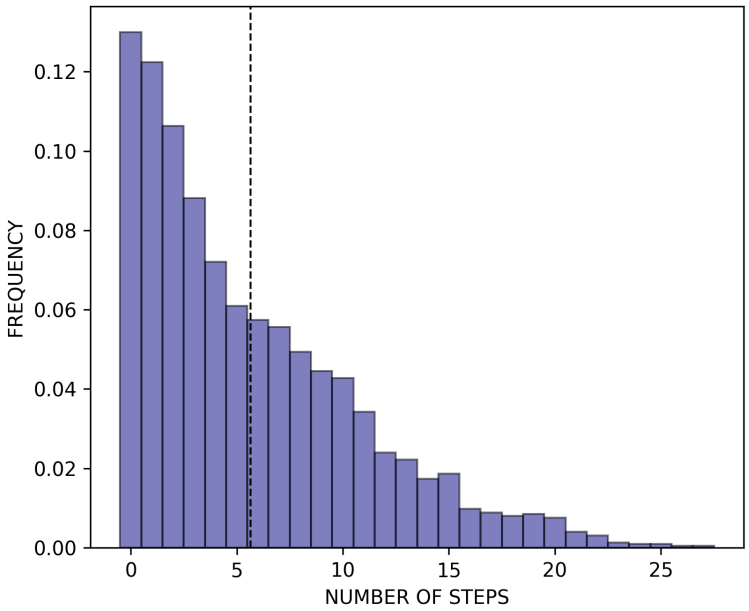

Figure 5.2. Probability distribution of the distance (number of steps) between the position where backtracking was started and the position where the problem became unsolvable. The dashed line shows the mean value. Bins are centered around integer values and have a width of 1. The fact that the distribution of distances is broad and (qualitatively) decreases slowly from 0, suggests that the simple unsolvability of the task is not the only trigger for backtracking. This goes against the idea that a plan is continuously monitored and revised.

5.3.2 The determinants of backtracking: a taxonomy

Because of this result, we hypothesized that the backtracking could be triggered by different signals, and that the unsolvability of the task may not be the only one. We propose a taxonomy that might help differentiate among these signals (see subsection 5.2.3). The percentage of backtracking events that fall into each of the categories of the taxonomy is shown in 5.3.

In almost 16% of the cases the backtracking was done when it was not necessary.

In more than half of the cases the remaining cases, backtracking followed a visual feedback of the non-solvability of the task (such as the bisection of the map or the impossibility of reaching any other reward).

In the remaining cases (less than half) no immediate visual feedback was provided. However, we identified two other possible conditions that could have induced participants to replan and which are based upon some sort of short term prospection. This is an intermediate condition between the simple perceptual domain and the planning one, since in this case there is not an evaluation of possible futures. In particular in around 9 % of the cases the backtrack was done after the creation of a second dead-end containing a reward; in around 14 % of the cases it was done after the prospected proximal bisection. The remaining unsolvable cases (around 17% of the total cases) contain the cases where no obvious error signal had been provided. In these cases, and also in the cases where the backtracking was not necessary, the participants might have been monitoring of the plan and subsequently having decided to revise it (though not correctly in all the cases).

The cases of proximal bisection and bisection (respectively 15 % and 11 %) were separated into two different categories even though strictly correlated. The remaining 19% of the cases are cases where the participant cannot reach any of the remaining rewards.

This sort of distribution tells us that in many cases participants might be relying upon the visual feedback of an error, rather than a (possibly costly) continual monitoring.

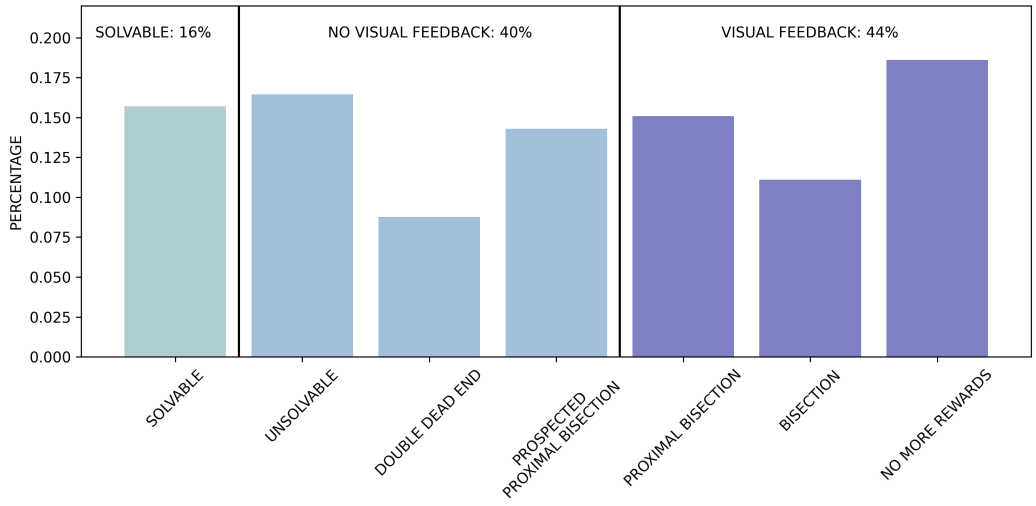

Figure 5.3. Taxonomy of the backtracking events. The figure shows the percentage of backtracking events that fall into each of the categories of the taxonomy. The taxonomy identifies three main groups: *No error*, *No visual feedback*, *Visual Feedback*. The first group includes backtracking events that were not necessary, the second group includes categories whose error detection required some sort of planning, short-term prospection or not immediate reasoning (which are: unsolvable, double dead-end, prospected proximal bisection). The group includes categories whose error detection was triggered by a visual feedback of the non-solvability of the task (which are: proximal bisection, bisection and no more rewards categories. See main text for the description).

5.4 Discussion

In this work we have shown that in the ThinkAhead task, people usually do not interrupt their plan immediately after the task has become unsolvable. In fact, the distribution of the distance between the position where backtracking was started and the position where the problem became unsolvable is broad and decreases slowly from 0. This suggests that the simple unsolvability of the task is not the only trigger for backtracking. This result helps narrow down the possible models to be considered to study planning in this task.

We proposed a taxonomy to help explain what might be the triggers for backtracking distinguishing between cases where the error detection was triggered by a visual feedback of the non-solvability of the task and cases where the error detection required some sort of planning, short-term prospection or not immediate reasoning.

A strong limitation of this taxonomy, is that they only have a descriptive power of the situation *where* the backtracking was triggered. However, if the participants actually backtracked *because* of that condition rather than something different is extremely difficult to prove. An interesting future work could be looking for what

are the dimensions created by the gaze and the hand kinematic that might help to detect in an unsuperivised way the backtracking events.
Another limitation of this study is that in the analysis are included all the cases where the backtracking was fixing an error due to motor control problems. These cases would require an ad-hoc detection strategy based on the rapidity of the error correction.

Chapter 6

Conclusions

Planning is an integral part of human cognition, manifesting in everyday tasks and complex problem-solving scenarios. Despite its ubiquity, studying planning is challenging due to its covert nature. This thesis aims to do this by combining multiple approaches: computational behavioral modelling, kinematic analysis, and eye-tracking measures, in a novel task named the ThinkAhead.

In chapter 1 we briefly lay the historical background of research on planning. We discuss the fundamental aspects of planning as a cognitive process and provide what is a common definition, intended as the process of evaluation of future scenarios and choice of the appropriate actions. The chapter introduces various models and theories that conceptualize planning as a search process and discusses how the problem of a bounded cognition prompted the use of heuristics and adaptive strategies as methods to manage the cognitive load. It also outlines the challenges of studying planning, such as the difficulty of capturing the hidden processes involved and the need for controlled experimental paradigms that rely on multiple measures.

Chapter 2 introduces the ThinkAhead task, a novel task that we designed to study planning in a controlled way. The task consists on navigating a 2D grid to connect nodes without revisiting any node, thus requiring substantial planning. The chapter details the experimental setup, where data were collected both online and in-person, including gaze and mouse movement recordings. The crucial property is that the task's design allows for varying the planning depth required to solve a problem and makes it suitable for analyzing different aspects of planning.

Chapter 3 investigates the adaptability of human planning strategies in the ThinkAhead task, specifically focusing on how individuals adjust their planning depth in response to varying task demands. The central question addressed is

whether participants employ a fixed planning depth across different problems or if they adaptively adjust their planning depth to optimize performance. Data were collected from 160 participants, but the analysis focused on 65 participants who completed at least 80 of the 90 problems. The study utilized a combination of logistic and linear regressions to analyze participants' success probability, problem completion time, and the total number of backtracks executed during the experiment.

The statistical analysis revealed significant effects of both the experimental design (level and depth) on success probability and problem completion time. Participants' success rates decreased as the complexity of the problems increased, highlighting the impact of problem level on performance. However, the success rates remained relatively stable across different levels when the planning depth was 1, indicating that simpler problems were consistently solved more successfully. The completion time analysis showed that participants took longer to solve problems at higher levels and greater planning depths. This suggests that more complex problems, requiring deeper planning, naturally took more time to solve. The interaction between problem level and planning depth was significant, indicating that the difficulty of the task was influenced by both the spatial complexity of the map and the cognitive demands of planning. The analysis of backtracking behavior provided additional insights into the planning process. While the total number of backtracks increased with problem complexity, the effect was not independent for the two dimensions (level and depth). This indicates that participants were more likely to revise their plans when faced with larger and more complex maps, but the propensity to backtrack was also influenced by the specific planning demands of the problem.

The model compared human performance to that of artificial agents with fixed and adaptive planning depths. The agents with adaptive planning depth adjusted their depth based on the minimum depth required for each problem, while fixed depth agents used the same depth for all problems.

One of the key findings of this chapter is the evidence of adaptive planning depth among participants. The comparison between human behavior and the computational models showed that participants adjusted their planning depth to match the minimum required depth for each problem.

The analysis of the initial plans (before the first backtrack) indicated that participants tended to start with a planning depth that was sufficient to solve the problem but not excessively deep. This adaptive strategy suggests a resource-rational approach, where individuals optimize their cognitive resources by selecting an appropriate planning depth that balances the need for accuracy and the cost of planning.

To ensure the robustness of these findings, control simulations were conducted. One such simulation involved a planner that mixed forward and backward strategy, which had prior knowledge of the last gem to be collected. This control aimed to test whether participants' planning depth was influenced by obvious cues in the problem layout. The results confirmed that the adaptive planning depth model still provided a better fit to participants' behavior, even when accounting for such cues. The findings from this chapter support the view that human problem-solving is a bounded rational process, where individuals adapt their cognitive resources to meet task demands. The use of flexible planning depth highlights an efficient strategy to manage the cognitive load associated with complex planning tasks.
The chapter also discusses the implications of these findings for understanding human cognition. The ability to adapt planning depth suggests a sophisticated level of metacognitive control, where individuals not only plan their actions but also plan how to plan. This metaplanning ability allows for dynamic adjustment of cognitive strategies based on the specific requirements of the task at hand.
Future research could explore the mechanisms underlying this adaptive planning behavior. Investigating how individuals determine the appropriate planning depth and the cues they use to make these decisions could provide deeper insights into the cognitive processes involved. Additionally, extending this research to dynamic and partially observable environments would test the generalizability of these findings and further our understanding of human adaptability in planning.

Chapter 4 explores the kinematics of eye-hand coordination during the ThinkAhead task to identify behavioral signatures of planning. The study investigates how the interplay between eye movements and hand movements during task execution reflects underlying planning processes.
The experiment was conducted in-person with participants using a computer and a mouse controller to solve the ThinkAhead task, while their gaze was tracked using an eye-tracker.
The analysis centered on identifying patterns of eye-hand coordination, particularly looking for instances of coarticulation. Coarticulation refers to the phenomenon where the execution of one action is influenced by subsequent actions, indicative of a planning processes.

The study found that participants exhibited a robust structure of sequential motor plans during task execution. These plans consisted of a fixation on a target followed by a reaching movement towards the target with the cursor. This alternating pattern allowed us to decompose the path implementation into discrete motor plans.
Within and between these motor plans, significant coarticulation effects were de-

tected. The analysis showed that the kinematics of mouse movements and fixation positions were influenced by future actions, suggesting that participants were not merely reacting to immediate stimuli but were continuously planning ahead.
To formally test the presence of planning, a linear classifier was trained to predict the next action based on segments of mouse trajectories and gaze positions. The classifier's accuracy significantly outperformed chance, demonstrating that current actions carried information about future movements. This predictive capability provided strong evidence for continuous planning processes in the ThinkAhead task.
The study also examined whether gaze behavior during fixations carried information about the direction of the next fixation beyond chance levels, and beyond models based on purely visual information (i.e.,no planning). By comparing accuracies for these models on the same classification test we showed that it was indeed the case.

The findings from this chapter highlight the intricate coordination between vision and action in planning. The presence of coarticulation effects indicates that planning is a dynamic and continuous process, where actions are interdependent and influenced by future goals. This challenges the traditional view of planning as a discrete and sequential process, instead proposing a model where planning and execution are closely intertwined.
The use of kinematic data provides a deeper understanding of the planning process, revealing aspects that are not apparent from final choices alone. The ability to predict future actions from current movements highlights the sophistication of human planning strategies and the importance of considering continuous dynamics in cognitive models. Further studies could also explore individual differences in planning strategies, examining how factors such as expertise, cognitive load, and task familiarity influence eye-hand coordination and planning efficiency. Understanding these individual variations would enhance the applicability of these insights across diverse populations and contexts.

Chapter 5 explores the participants' commitment to their plans and the triggers for backtracking. It was observed that backtracking often occurred after visual feedback indicated the task was unsolvable, suggesting that participants did not frequently replan but relied on perceptual cues. A taxonomy was proposed to differentiate the signals that trigger backtracking, providing insights into the interplay between planning and perceptual feedback.

This thesis represented an attempt to study human planning strategies through the innovative ThinkAhead task and a combination of multiple strategies. The findings underscore the adaptive nature of planning, where individuals modulate their

planning depth based on task complexity. The use of kinematic and eye-tracking data reveals the intricate coordination between vision and action.
Potential applications of this research might make use of the relations between eye movements and hand kinematics to predict user actions, enabling the development of interfaces that dynamically adjust the visual output to the expected users need, to improve efficiency and user experience.
Future research could extend these findings by exploring dynamic and partially observable environments, where task elements change over time or are initially unknown. Investigating how individuals adapt their planning strategies in such contexts would provide deeper insights into the flexibility and robustness of human planning. Additionally, examining the learning dynamics and the impact of different training sequences on planning could further elucidate how planning skills develop and improve over time.

Appendix A

Supplementary materials

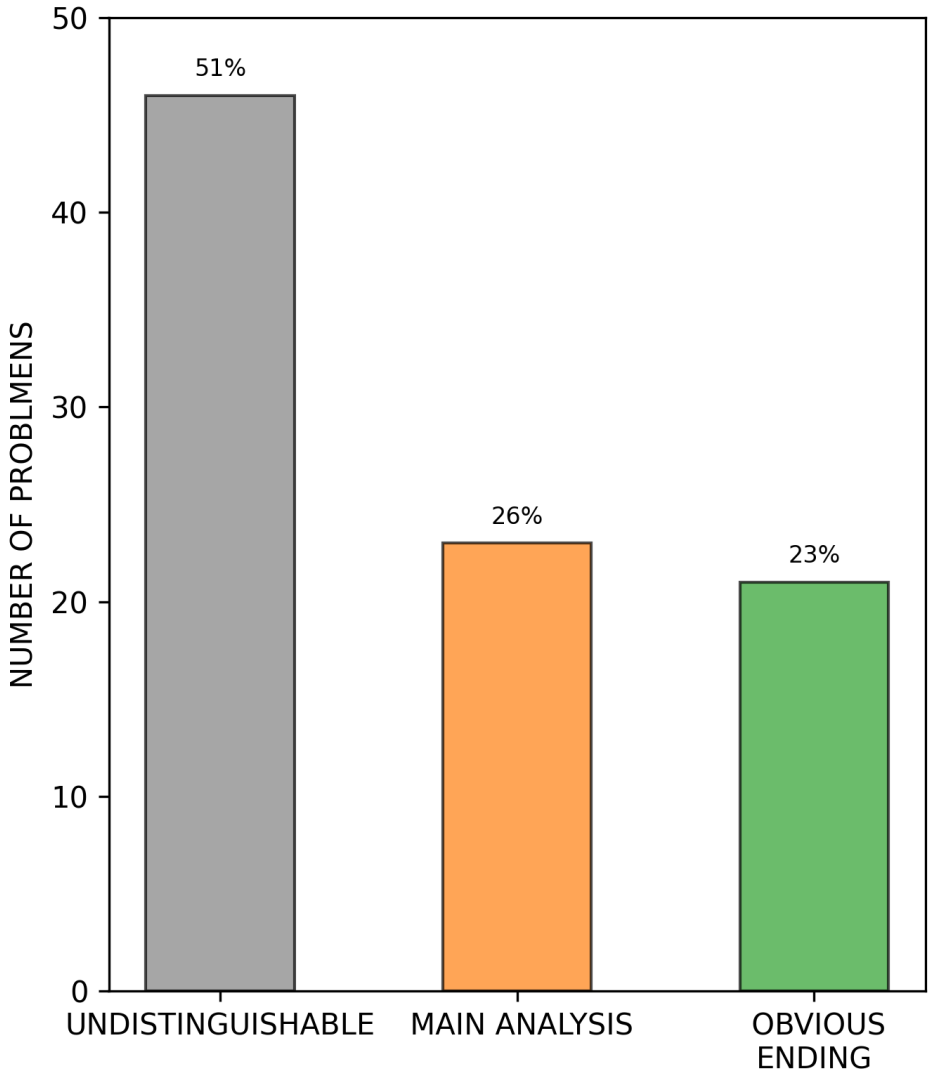

Figure A.1. Comparison of the performance of the planner used in the main analysis and of the "obvious ending" planner. About half of the problems (51%, n = 46) do not have any obvious ending and hence cannot distinguish between the two models. The planner used in the main analysis and the "obvious ending" planner better explain participants' behaviour in 26% (n = 23) and 23% (n = 21) of the remaining problems, respectively.

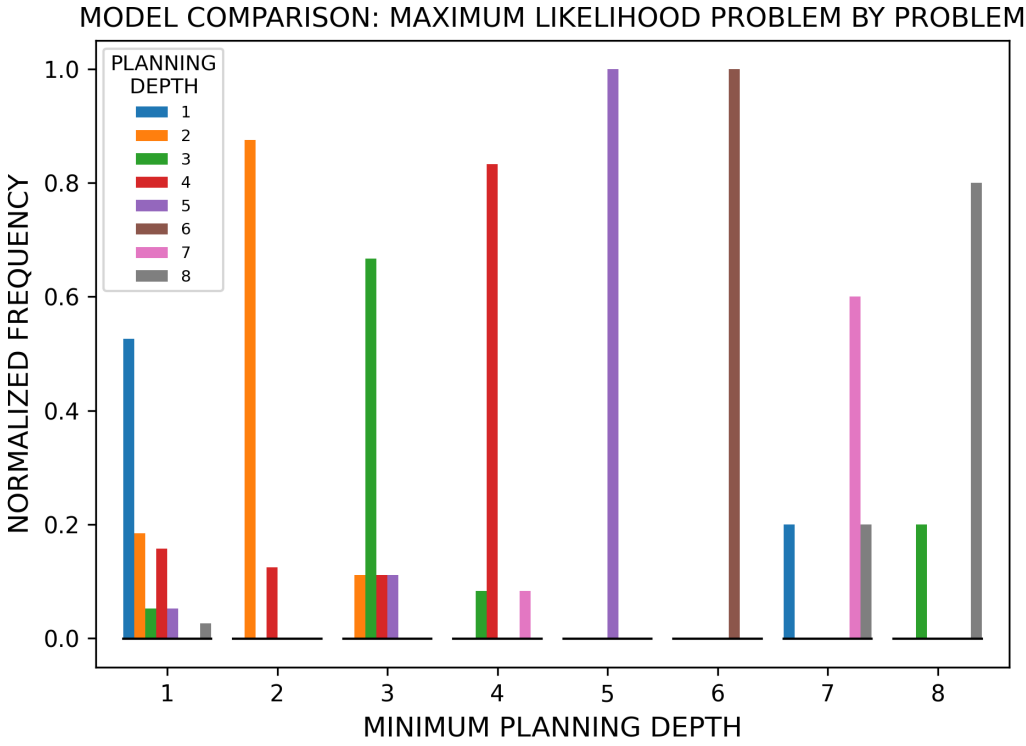

Figure A.2. Comparison between participants and planning models, for the "obvious ending" planner. The figure shows the number of times that each "obvious ending" planning model from depth 1 to 8 had the maximum likelihood of the gems collected before the first backtrack by the participants, in the problems 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).

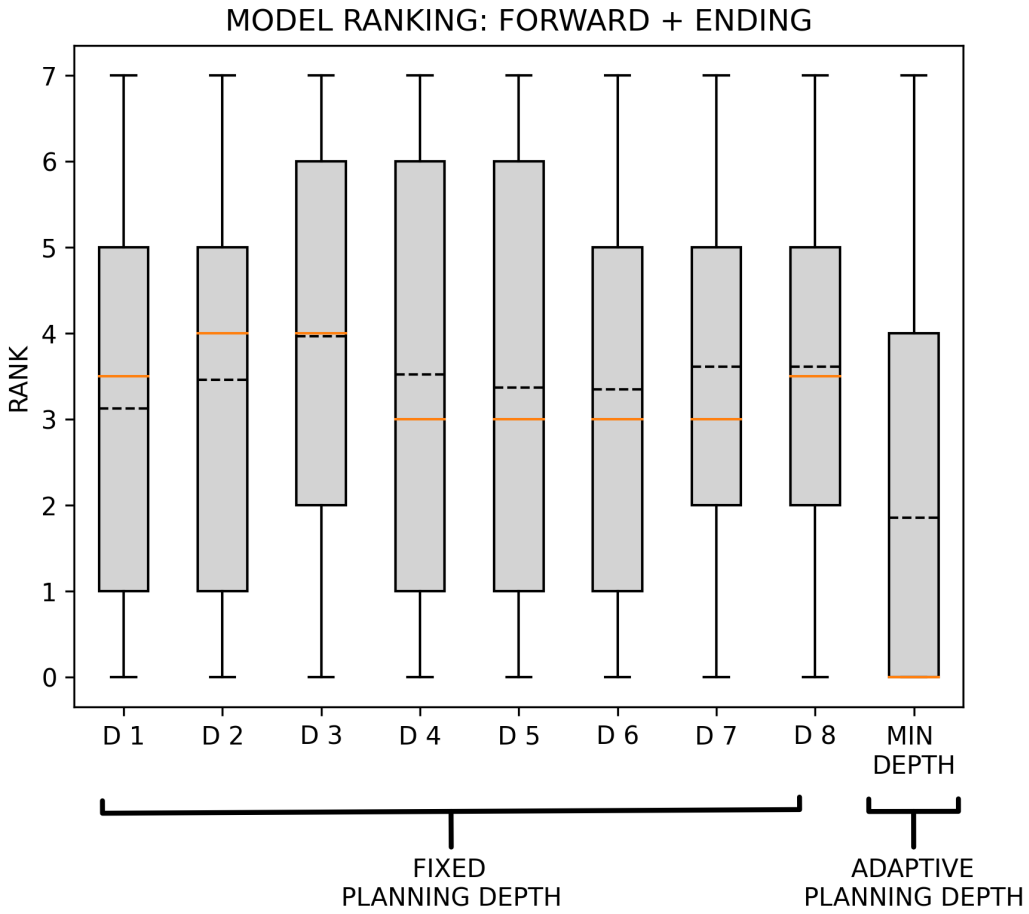

Figure A.3. Rank values across the 90 problems, for the "obvious ending" planner with fixed or adaptive planning depth. The figure shows the distributions of ranks of planners using a fixed planning depth (D1-D8) across all the problems and of the planner using a planning depth adapted to the minimum 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. The adaptive planning depth distribution is significantly smaller than any of the other distributions ($\chi^2(1, 90) = 31.51, p - value < 5.0 * 10 - 5$).

Velocity threshold detection

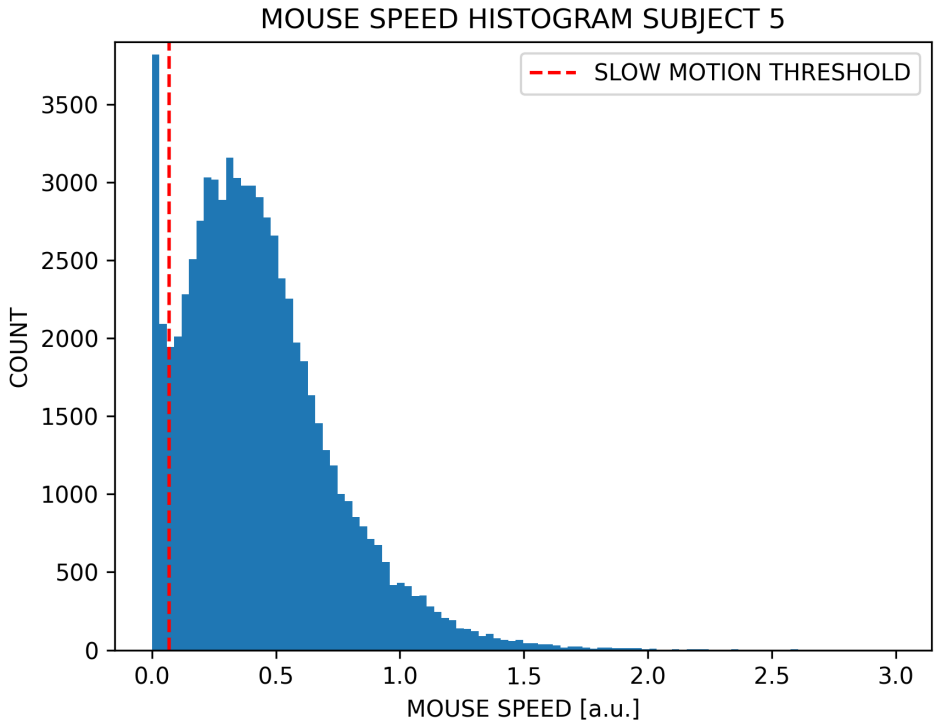

Figure A.4. Velocity threshold detection. The figure shows the distribution of the velocity of the mouse cursor in the ThinkAhead task. The red line shows the threshold used to detect the end of a movement.

By looking at the distribution of the velocity of the mouse cursor in the ThinkAhead task (Figure A.4), we can see that the velocity of the mouse cursor is bimodal. One peak corresponds to the (higher) velocity of the mouse cursor when the participant is moving the mouse cursor, while the other peak corresponds to the (lower) velocity of the mouse cursor when the participant is not moving the mouse cursor or is doing it very slowly. The threshold used to detect the end of a movement is set at the minimum between the two peaks. This threshold was thus computed for each participant.

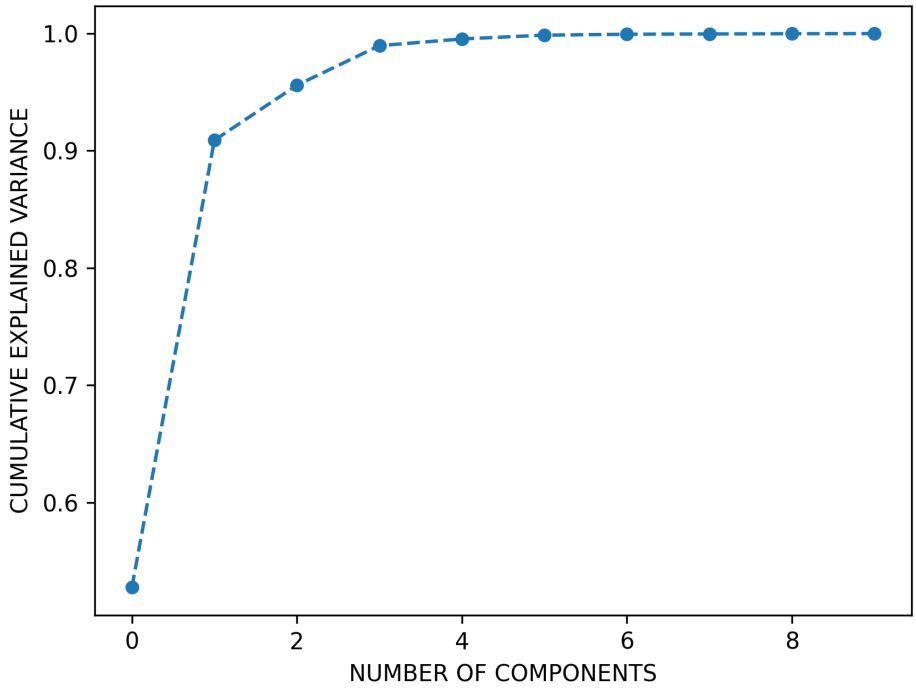

Figure A.5. The figure shows the explained variance of the principal components of the mouse trajectory data

Bibliography

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

Ames, K. C., Ryu, S. I., & Shenoy, K. V. (2019). Simultaneous motor preparation and execution in a last-moment reach correction task. *Nature Communications*, *10*, 2718. https://doi.org/10.1038/s41467-019-10772-2

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

Azadi, R., Zhu, E. Y., & McPeek, R. M. (2021). Modulation of saccade trajectories during sequential saccades [PMID: 33471606]. *Journal of Neurophysiology*, *125*(3), 796–804. https://doi.org/10.1152/jn.00106.2020

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

Baldauf, D., & Deubel, H. (2008). Properties of attentional selection during the preparation of sequential saccades. *Experimental Brain Research*, *184*, 411–425. https://doi.org/10.1007/s00221-007-1114-x

Barca, L., & Pezzulo, G. (2012). Unfolding visual lexical decision in time. *PloS one*, *7*(4), e35932.

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

Binder, F. J., Mattar, M. G., Kirsh, D., & Fan, J. E. (2023). Humans choose visual subgoals to reduce cognitive cost. *Proceedings of the Annual Meeting of the Cognitive Science Society*, *45*. https://escholarship.org/uc/item/74q3976x

Botvinick, M., & Toussaint, M. (2012). Planning as inference. *Trends in Cognitive Sciences*, *16*(10), 485–488. https://doi.org/https://doi.org/10.1016/j.tics.2012.08.006

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

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., et al. (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.

Crowe, D. A., Averbeck, B. B., Chafee, M. V., Anderson, J. H., & Georgopoulos, A. P. (2000). Mental maze solving. *Journal of Cognitive Neuroscience*, *12*(5), 813–827. https://doi.org/10.1162/089892900562426

Daniloff, R., & Hammarberg, R. (1973). On defining coarticulation. *Journal of Phonetics*, *1*(3), 239–248. https://doi.org/https://doi.org/10.1016/S0095-4470(19)31388-9

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*. De Gruyter Mouton.

De Vries, J. P., Hooge, I. T., & Verstraten, F. A. (2014). Saccades toward the target are planned as sequences rather than as single steps. *Psychological Science*, *25*(1), 215–223. https://doi.org/10.1177/0956797613497020

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

Fiedler, S., & Glöckner, A. (2012). The dynamics of decision making in risky choice: An eye-tracking analysis. *Frontiers in Psychology*, *3*. https://doi.org/10.3389/fpsyg.2012.00335

Fowler, C. A., & Saltzman, E. (1993). Coordination and coarticulation in speech production. *Language and Speech*, *36*(2-3), 171–195. https://doi.org/10.1177/002383099303600304

Freeman, J. B., Dale, R., & Farmer, T. A. (2011). Hand in motion reveals mind in motion. *Frontiers in psychology*, *2*, 59.

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

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

Fudenberg, D., Newey, W., Strack, P., & Strzalecki, T. (2019). Testing the drift-diffusion model. *Proceedings of the National Academy of Sciences*, *117*, 33141–33148. https://doi.org/10.1073/pnas.2011446117

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. (1979a). A cognitive model of planning. *Cogn. Sci.*, *3*, 275–310. https://doi.org/10.1207/s15516709cog0304_1

Hayes-Roth, B., & Hayes-Roth, F. (1979b). A cognitive model of planning. *Cognitive Science*, *3*(4), 275–310.

Hayhoe, M., & Ballard, D. (2005). Eye movements in natural behavior. *Trends in Cognitive Sciences*, *9*(4), 188–194. https://doi.org/10.1016/j.tics.2005.02.009

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.

Huang, J., Velarde, I., Ma, W. J., & Baldassano, C. (2023). Schema-based predictive eye movements support sequential memory encoding. *eLife*, *12*, e82599. https://doi.org/10.7554/eLife.82599

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. (2014). Computational psychiatry. https://doi.org/10.1007/978-1-4614-7320-6_501-2

Huys, Q., Browning, M., Paulus, M., & Frank, M. (2020). Advances in the computational understanding of mental illness. *Neuropsychopharmacology*, *46*, 3–19. https://doi.org/10.1038/s41386-020-0746-4

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

Huys, Q. J., 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

Itti, L., & Koch, C. (2001). Computational modeling of visual attention. *Nature Reviews Neuroscience*, *2*, 194–203.

Jammalamadaka, S. R., & SenGupta, A. (2001). *Topics in circular statistics* [Section 8.3]. World Scientific Press.

Jensen, K. T., Hennequin, G., & Mattar, M. G. (2023). A recurrent network model of planning explains hippocampal replay and human behavior. *bioRxiv*. https://www.biorxiv.org/content/10.1101/2023.01.17.426598v1

Jerde, T. E., Soechting, J. F., & Flanders, M. (2003). Coarticulation in fluent fingerspelling. *Journal of Neuroscience*, *23*(6), 2383–2393.

Kafadar, E., Fisher, V. L., Quagan, B., Hammer, A., Jaeger, H., Mourgues, C., Thomas, R., Chen, L., Imtiaz, A., Sibarium, E., Negreira, A. M., Sarisik, E., Polisetty, V., Benrimoh, D., Sheldon, A. D., Lim, C., Mathys, C., & Powers, A. R. (2022). Conditioned hallucinations and prior overweighting are state-sensitive markers of hallucination susceptibility [Cortical Mechanisms of Psychosis]. *Biological Psychiatry*, *92*(10), 772–780. https://doi.org/https://doi.org/10.1016/j.biopsych.2022.05.007

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

Kirsh, D., & Maglio, P. (1994). On distinguishing epistemic from pragmatic action. *Cognitive Science*, *18*(4), 513–549. https://doi.org/https://doi.org/10.1207/s15516709cog1804_1

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. (1987). Planning as search: A quantitative approach. *Artificial Intelligence*, *33*(1), 65–88.

Krotov, A., Russo, M., Nah, M., Hogan, N., & Sternad, D. (2022). Motor control beyond reach—how humans hit a target with a whip. *R. Soc. Open Sci.*, *9*, 220581.

Krusche, M. J. F., Schulz, E., Guez, A., & Speekenbrink, M. (2018). Adaptive planning in human search. *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.

Land, M. (2004). Eye movements in daily life. In L. M. Chalupa & J. S. Werner (Eds.), *The visual neurosciences* (pp. 1357–1368, Vol. 2). MIT Press.

Land, M., & McLeod, P. (2000). From eye movements to actions: How batsmen hit the ball. *Nature Neuroscience*, *3*, 1340–1345. https://doi.org/10.1038/81887

Land, M. F., Mennie, N., & Rusted, J. (1999). Eye movements and the roles of vision in activities of daily living: Making a cup of tea. *Perception*, *28*, 1311–1328.

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.

MacKay, D. J. (2003). *Information theory, inference and learning algorithms*. Cambridge university press.

Maisto, D., Donnarumma, F., & Pezzulo, G. (2015). divide et impera: Subgoaling reduces the complexity of probabilistic inference and problem solving. *Journal of The Royal Society Interface*, *12*(104), 20141335. https://doi.org/10.1098/rsif.2014.1335

Maisto, D., Donnarumma, F., & Pezzulo, G. (2016). Nonparametric problem-space clustering: Learning efficient codes for cognitive control tasks [[IF 2015: 1.743]]. *Entropy*, *18*(2), 61.

Mannan, S. K., Ruddock, K. H., & Wooding, D. S. (1997). Fixation patterns made during brief examination of two-dimensional images. *Perception*, *26*, 1059–1072.

Maselli, A., De Pasquale, P., Lacquaniti, F., & d'Avella, A. (2022). Interception of virtual throws reveals predictive skills based on the visual processing of throwing kinematics. *iScience*, *25*(10), 105212. https://doi.org/https://doi.org/10.1016/j.isci.2022.105212

Maselli, A., Dhawan, A., Cesqui, B., Russo, M., Lacquaniti, F., & D'Avella, A. (2017). Where are you throwing the ball? i better watch your body, not just your arm! *Frontiers in Human Neuroscience*, *11*. https://doi.org/10.3389/fnhum.2017.00505

Maselli, A., Dhawan, A., Russo, M., Cesqui, B., Lacquaniti, F., & d'Avella, A. (2019). A whole body characterization of individual strategies, gender differences, and common styles in overarm throwing [PMID: 31577474]. *Journal of Neurophysiology*, *122*(6), 2486–2503. https://doi.org/10.1152/jn.00011.2019

Maselli, A., Gordon, J., Eluchans, M., Lancia, G. L., Thiery, T., Moretti, R., Cisek, P., & Pezzulo, G. (2023). Beyond simple laboratory studies: Developing sophisticated models to study rich behavior. *Physics of Life Reviews*.

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

McDermott, D., & Davis, E. (1983). Planning routes through uncertain territory. *Artif. Intell.*, *22*, 107–156. https://doi.org/10.1016/0004-3702(84)90045-6

McDermott, D. (1978). Planning and acting*. *Cognitive Science*, *2*(2), 71–100. https://doi.org/https://doi.org/10.1207/s15516709cog0202_1

McPeek, R. M., Skavenski, A. A., & Nakayama, K. (2000). Concurrent processing of saccades in visual search. *Vision Research*, *40*(18), 2499–2516. https://doi.org/https://doi.org/10.1016/S0042-6989(00)00102-4

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.

Montague, P., Dolan, R., Friston, K. J., & Dayan, P. (2012). Computational psychiatry. *Trends in Cognitive Sciences*, *16*, 72–80. https://doi.org/10.1016/j.tics.2011.11.018

Myers, C., Interian, A., & Moustafa, A. (2022). A practical introduction to using the drift diffusion model of decision-making in cognitive psychology, neuroscience, and health sciences. *Frontiers in Psychology*, *13*. https://doi.org/10.3389/fpsyg.2022.1039172

Najemnik, J., & Geisler, W. S. (2005). Optimal eye movement strategies in visual search. *Nature*, *434*(7031), 387–391. https://doi.org/10.1038/nature03390

Newell, A., Simon, H. A., et al. (1972). *Human problem solving* (Vol. 104). Prentice-hall Englewood Cliffs, NJ.

Papadimitriou, C. H., & Yannakakis, M. (1991). Shortest paths without a map. *Theoretical Computer Science*, *84*(1), 127–150.

Parkhurst, D., Law, E., & Niebur, E. (2002). Modeling the role of salience in the allocation of overt visual attention. *Vision Research*, *42*, 107–123.

Parkhurst, D. J., & Niebur, E. (2003). Scene content selected by active vision. *Spatial Vision*, *16*(2), 125–154. https://doi.org/10.1163/15685680360511645

Parr, T., Pezzulo, G., & Friston, K. J. (2022). *Active inference: The free energy principle in mind, brain, and behavior*. MIT Press.

Pelz, J. B., & Canosa, R. (2001). Oculomotor behavior and perceptual strategies in complex tasks. *Vision Research*, *41*(25), 3587–3596. https://doi.org/https://doi.org/10.1016/S0042-6989(01)00245-0

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

Pezzulo, G., Rigoli, F., & Friston, K. (2018). Hierarchical active inference: A theory of motivated control. *Trends in Cognitive Sciences*, *22*(4), 294–306.

Radach, R., Heller, D., & Inhoff, A. (1999). Occurrence and function of very short fixation durations in reading, 321–331. https://doi.org/10.1007/978-1-4757-3054-8_46

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological review*, *85*(2), 59.

Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual search. *Quarterly Journal of Experimental Psychology*, *62*(8), 1457–1506. https://doi.org/10.1080/17470210902816461

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.

Robinson, D. A. (1963). A method of measuring eye movemnent using a scieral search coil in a magnetic field. *IEEE Transactions on Bio-medical Electronics*, *10*(4), 137–145. https://doi.org/10.1109/TBMEL.1963.4322822

Rothkopf, C. A., Ballard, D. H., & Hayhoe, M. M. (2016). Task and context determine where you look. *Journal of Vision*, *7*(14), 16. https://doi.org/10.1167/7.14.16

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. *Psychological Science*.

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.

Sacerdoti, E. D. (1974). Planning in a hierarchy of abstraction spaces. *Artificial Intelligence*, *5*(2), 115–135. https://doi.org/https://doi.org/10.1016/0004-3702(74)90026-5

Sacerdoti, E. D. (1975). *A structure for plans and behavior* (Technical Report). SRI International, Artificial Intelligence Center. Menlo Park, CA.

Shah, A., Barto, A. G., & Fagg, A. H. (2013). A dual process account of coarticulation in motor skill acquisition. *Journal of motor behavior*, *45*(6), 531–549.

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

Shannon, C. E. (1988). Programming a computer for playing chess. In D. Levy (Ed.), *Computer chess compendium* (pp. 2–13). Springer New York. https://doi.org/10.1007/978-1-4757-1968-0_1

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; 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., 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

Song, J.-H., & Nakayama, K. (2009). Hidden cognitive states revealed in choice reaching tasks. *Trends in cognitive sciences*, *13*(8), 360–366.

Stone, M. (1960). Models for choice-reaction time. *Psychometrika*, *25*(3), 251–260.

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

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 Dijk, S. G., & Polani, D. (2011). Grounding subgoals in information transitions. *2011 IEEE Symposium on Adaptive Dynamic Programming and Reinforcement Learning (ADPRL)*, 105–111. https://doi.org/10.1109/ADPRL.2011.5967384

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

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.

Whalen, D. (1990). Coarticulation is largely planned. *Journal of Phonetics*, *18*(1), 3–35. https://doi.org/https://doi.org/10.1016/S0095-4470(19)30356-0

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.

Yarbus, A. L. (1967). *Eye movements and vision*. Plenum Press. https://doi.org/10.1007/978-1-4899-5379-7

Zhao, M., & Marquez, A. G. (2013). Understanding humans' strategies in maze solving. *arXiv preprint arXiv:1307.5713*.

Zhu, S., Lakshminarasimhan, K. J., Arfaei, N., & Angelaki, D. E. (2022). Eye movements reveal spatiotemporal dynamics of visually-informed planning in navigation (H. Zhang, J. I. Gold, & H. J. Spiers, Eds.). *eLife*, *11*, e73097. https://doi.org/10.7554/eLife.73097