Statistical Properties of the Environment Influence Action Abstraction in Human Planning

Valkyrie Felso[1], **Maximilian Mittenbühler**[2] **& Ruiqi He**[1]
[1]Max Planck Institute for Intelligent Systems, Tübingen, Germany
[2]Department of Computer Science, University of Tübingen, Tübingen, Germany

Abstract

Planning requires significant cognitive effort, as it involves constructing and evaluating numerous potential futures. A key method to simplify planning is the grouping of sequences of actions into cohesive units—often referred to as action chunking. While studies confirm the existence of action chunking in human planning, we explore chunking as a learning process across multiple tasks with different statistical properties. To that end, we ran a large-scale preregistered study using the Lightbot navigation paradigm, in which we manipulated the underlying statistical properties of the planning tasks. We found that the majority of participants adapted their chunking behaviour to match the environment's statistical properties by either learning to use the action chunk that is optimal given these statistical properties or a subsequence thereof. Together, our results suggest that statistical properties of the environment influence action chunking in human planning.

Keywords: planning; action abstraction; resource-rational

Introduction

Planning is a computationally demanding process that requires selecting and sequencing actions to achieve future goals. Despite the combinatorial complexity of many planning problems, humans are often able to plan effectively in real-world environments. Previous work suggests that this apparent efficiency stems from people simplifying planning problems in systematic ways rather than solving them in a fully detailed action space (Ho et al., 2022; Niv, 2019; Tomov et al., 2020). One powerful form of simplification is *action chunking*, also known as action abstraction (Correa et al., 2025; Dezfouli & Balleine, 2013; Éltető & Dayan, 2023), which describes the grouping of multiple low-level, atomic actions into a single higher-level unit that can be treated as a single step or a single action during planning. For example, when cutting a tomato, one does not explicitly plan each individual motor action — such as raising the knife, positioning it, applying force, and retracting it — but instead treats this sequence as a single coherent operation of "cutting the tomato". When planning how to cook dinner, it is therefore much simpler to just consider this single action chunk instead of each atomic action. Action chunking plays a fundamental role in planning frameworks, such as hierarchical reinforcement learning (Botvinick et al., 2009; Eckstein & Collins, 2021; Sutton et al., 1999) and has been implicated in effective generalisation of behaviour (Badre et al., 2010).

However, how beneficial a particular action chunk is for planning depends on multiple factors. On the one hand, chunking reduces cognitive demands by simplifying the planning problem; on the other hand, chunking reduces flexibility, as committing to a chunk limits one's ability to adapt mid-sequence. This trade-off can be formalised within a resource-rational framework (Lieder & Griffiths, 2020), which characterises the usefulness of a chunk given both environmental rewards and cognitive constraints. Whether a chunk is resource-rational therefore depends on the statistics of the environment, such as how likely certain action transitions are or how often they occur. In this work, we use the term "resource-rational" to denote the action chunk that maximises expected utility given the environmental statistics and the costs associated with planning.

Previous research has provided strong evidence that people form and use action chunks during planning (Correa et al., 2025; Dezfouli & Balleine, 2013). However, much less is known about how people learn which chunks to use, and in particular how statistical properties of the environment influence people's chunking behaviour. Humans are known to exploit statistical regularities through repeated interaction with their environment (Solway et al., 2014), suggesting that chunk formation may somehow be sensitive to statistical properties. Therefore, in this work, we investigate how two concrete statistical properties of the environment influence people's action chunking: First, we manipulate the length of the resource-rational action chunk, comparing environments in which it is optimal to use chunks consisting of three or four actions. Second, we manipulate how certain a chunk is, operationalised as the transition probabilities between the atomic actions that compose the chunk. Higher transition probabilities indicate that the actions within a chunk reliably occur together, making chunking more likely to be advantageous, whereas lower transition probabilities reduce the expected utility of a chunk. Finally, we manipulate the order of low and high transition probabilities within a chunk. If high transition probabilities occur late, it would be more resource-rational to use a longer chunk. If high transition probabilities occur early in a chunk, it would be more resource-rational to use a shorter chunk.

To examine how these factors influence human chunking behaviour, we employ a planning task in which resource-rational chunk length and transition probabilities varied systematically. This allows us to test four preregistered hypotheses (OSF: https://osf.io/a4zvy/overview?view_only=c6c735e956ef435da89645b596461602):

1. Participants' average chunk lengths increase in environments which favour longer action sequences.
2. Participants' average chunk lengths increase when longer chunks are more likely to be useful, as reflected by higher transition probabilities.
3. There exists an interaction effect such that transition probabilities exert a stronger influence on abstraction length for longer action sequences.
4. Environments in which transition probabilities increase toward the end of a sequence elicit longer chunks than environments in which transition probabilities decrease.

By observing the chunking behaviour of 363 participants across environments with different statistical structures, this work aims to shed light on how people adaptively simplify planning problems in a resource-rational manner.

Method

Modified Lightbot task

We adopted a modified version of the Lightbot navigation paradigm (Correa et al., 2025; Sanborn et al., 2018), in which participants guide a robot avatar through a maze to reach a goal by selecting and inputting appropriate action commands (see Figure 1). This task is a process-tracing paradigm, allowing researchers insight into how people construct action chunks. On each trial, participants input a list of commands to control the robot, choosing from four atomic actions: *walk*, *jump*, *turn left*, and *turn right*. The robot executes the commands in the order they appear in the *Main* box. The trial ends once the robot reaches the goal tile (shown in blue), upon which the robot will automatically light the blue tile and proceed to the next trial.

The robot has limited units of energy, which are consumed whenever it follows a command in the *Main* box. In addition to the four main actions, participants can also create one "memory" by dragging and dropping actions into the *Memory* box. *Memories* function as action chunks: when a *Memory* is reused in later trials, the entire sequence is executed at a reduced energy cost. The contents of the *Memory* box persist across trials unless explicitly modified by the participant. Whenever the *Memory* is created or changed, no energy savings are applied on that trial, reflecting the cognitive cost of constructing or revising a chunk. However, on subsequent trials in which the *Memory* list remains unchanged, reusing the stored chunk reduces the energy consumed by the robot. In this way, the task explicitly implements a trade-off between the short-term cost of creating or modifying a chunk and the long-term benefits of reusing it and operationalises the idea that action chunks reduce cognitive and energetic costs when reused, while also capturing the cost of committing to and maintaining a chunk.

To incentivise participants to discover and maintain chunks that efficiently solved the series of planning problems, they were awarded a monetary bonus proportional to the number of energy units saved. The current energy budget and

Figure 1: Three example trials of the Lightbot task where the participant tried out different chunks by adding and removing atomic actions in the memory box. All three shown trials are from the same map set and therefore all share the same resource-rational chunk of *jump*, *right*, *jump*, *walk* with transition probabilities 99% ⟶ 99% ⟶ 80%.

energy consumption were displayed prominently at the top of the screen, making the cost-benefit structure of chunking transparent.

Figure 1 illustrates three example trials from one hypothetical participant. In the first trial (Trial 6 of 30), the hypothetical participant reuses a previously stored chunk consisting of the actions (*jump*, *turn right*). By reusing this chunk, the participant saves one unit of energy, which is reflected in an increased bonus shown on the following trial (from $0.25 to $0.31). In the second trial (Trial 7 of 30), the hypothetical participant modifies the *Memory* list to include a longer chunk (*jump*, *turn right*, *jump*). Because the *Memory* was changed, no energy is saved on that trial, resulting in no bonus earned on that trial, as reflected in the upper right red and blue boxes where the instruction count equals the count of remaining energy. Finally, in the next trial (Trial 8 of 30), the participant reuses this new chunk without further modification, and saves two units of energy, resulting in additional bonus. In the next unshown trial (Trial 9 of 30), the bonus displayed in the upper left corner would reflect this.

Conditions

Participants were assigned to one of 60 granular conditions, which we call different "map sets". Each map in the map set comprised a sequence of 30 consecutive trials. Within each map set, all trials shared the same underlying resource-rational action abstraction, drawn from one of four possibilities: 1) *jump* → *jump* → *turn right*, 2) *jump* → *turn right* → *jump*, 3) *jump* → *jump* → *turn right* → *walk*, or 4) *jump* → *turn right* → *jump* → *walk*. That is, the same optimal chunk length and statistical structure made a particular chunk resource-rational across the 30 trials. This design encouraged participants to discover and exploit regularities shared across trials, rather than solving each trial independently. The 60 map sets were grouped into six experimental conditions based on their statistical properties. Conditions varied along two dimensions: (1) the length of the resource-rational action chunk (three versus four actions), and (2) the transition probabilities between the atomic actions composing the chunk (0.8 vs. 0.99) (see Table 1 for details on each condition). In other words, each condition instantiated one of the four action chunks that would solve the 30-trial sequence optimally in terms of energy expenditure. The transition probabilities modulated how reliably these action chunks yielded optimal solutions, thereby controlling the extent to which a given chunk was statistically favoured across trials.

Participants and procedure

Participants were recruited via Prolific (Palan & Schitter, 2018). Only participants fluent in English were eligible for the study, and we excluded all participants who had participated in pilot experiments due to familiarity with the task. Our final sample consisted of 534 participants (274 men, 256 women, 4 non-binary participants; 18–75 years old, mean 33.72 years old).

Table 1: Overview of the conditions and the corresponding transition probabilities, resource-rational chunk length and number of participants in each condition after exclusion.

Cond.	Transition probabilities	Chunk length	No. of participants
1	$a_1 \xrightarrow{0.8} a_2 \xrightarrow{0.99} a_3$	3	48
2	$a_1 \xrightarrow{0.99} a_2 \xrightarrow{0.99} a_3$	3	92
3	$a_1 \xrightarrow{0.99} a_2 \xrightarrow{0.8} a_3$	3	46
4	$a_1 \xrightarrow{0.8} a_2 \xrightarrow{0.99} a_3 \xrightarrow{0.99} a_4$	4	43
5	$a_1 \xrightarrow{0.99} a_2 \xrightarrow{0.99} a_3 \xrightarrow{0.99} a_4$	4	90
6	$a_1 \xrightarrow{0.99} a_2 \xrightarrow{0.99} a_3 \xrightarrow{0.8} a_4$	4	44

Participants were first presented with an extensive tutorial on the Lightbot task. The tutorial included a comprehension check focused on how energy savings translated into monetary bonus, ensuring that participants understood the incentive structure of the task. After the tutorial, participants were quizzed more generally on their understanding of the task. After successfully completing the tutorial and comprehension checks, participants completed 30 trials of the Lightbot task. Upon finishing the task, participants provided demographic information, reported any prior experience with the Lightbot task, and were given the opportunity to provide feedback or report technical issues. The experiment had a median completion time of 66.35 minutes. Participants received a base payment of $10, with performance-contingent bonuses averaging $1.25.

Results

According to our preregistered exclusion criteria, we excluded 171 participants who either never formed any action abstractions, demonstrated misunderstanding of the task, or exceeded the allowed number of quiz attempts (median bonus: $0.00). The final sample comprised 363 participants (median bonus: $1.77) distributed across the six conditions (see Table 1). Below we report results of non-parametric statistical tests due to non-normality (all $W < 0.82$, $p < .001$) and non-homogeneity ($W = 10.21$, $p < .001$) of the residuals for hypotheses 1–3 (as preregistered).

Effect of optimal action sequence length Hypothesis 1 predicted that participants would form longer action abstractions in environments where longer action chunks are resource-rational. To test this, we compared and grouped together environments in which an action sequence of length three was resource-rational (conditions 1, 2, 3) to those in which a sequence of length four was resource-rational (conditions 4, 5, 6). We found that participants' average chunk length in the last 15 trials was indeed longer when the resource-rational chunk length was four actions long compared to three actions long, as indicated by a significant one-sided Mann-Whitney U test (length 3: $\mathrm{M} = 2.58, \mathrm{SD} = 1.10$; length 4:

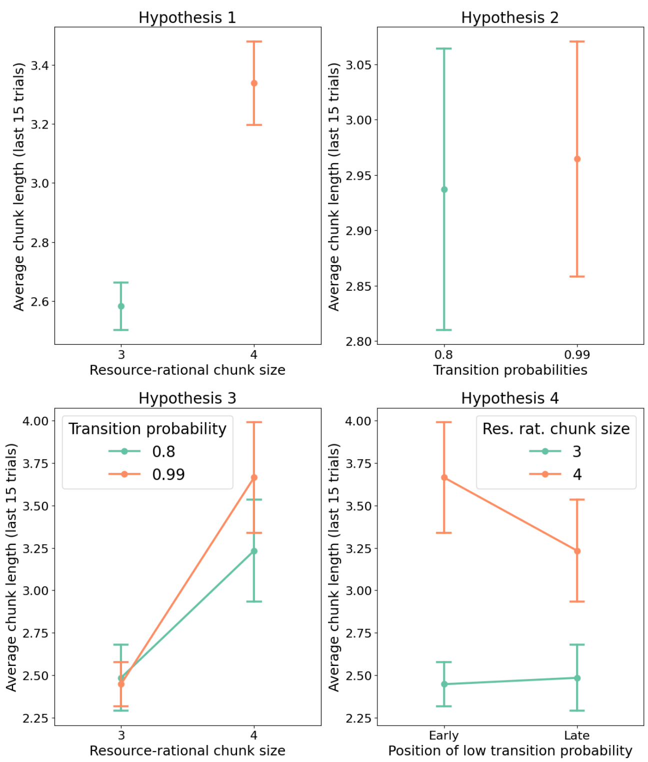

Figure 2: Each subfigure corresponds to one of our preregistered hypotheses. Error bars indicate standard error of the mean.

M = 3.34, SD = 1.88; U = 13313, $p < .001$), providing support for hypothesis 1 (see also the top left subfigure in Figure 2). This suggests that participants adjusted the chunk length in a resource-rational manner based on the optimal length of action chunks in the set of planning tasks.

Effect of transition certainty Hypothesis 2 predicted that abstraction length would increase when longer action sequences were more certainly useful, operationalised as higher transition probabilities between actions. We tested this by again grouping and comparing environments with consistently high transition probabilities (conditions 2 and 5) to environments with lower transition probabilities (conditions 1, 3, 4, 6). A one-sided Mann-Whitney U test comparing the average chunk lengths between the two groups yielded a non-significant result (high certainty: M = 2.96, SD = 1.43; low certainty: M = 2.94, SD = 1.71; U = 17886.50, $p = .071$), failing to support hypothesis 2 (see the top right subfigure in Figure 2). Thus, transition certainty had no influence on abstraction length.

Interaction effect of transition certainty with sequence length Hypothesis 3 predicted an interaction effect such that transition probabilities would exert a stronger influence on abstraction length for longer action sequences. The preregistered Kruskal-Wallis test was significant ($H = 16.86, p < .001$). Bonferroni-corrected post-hoc Dunn's tests revealed a significant difference in medians from the group with low certainty and a resource-rational chunk length three (M = 2.46, SD = 1.12), particularly between the group with low certainty and a resource-rational chunk length four (M = 3.45, SD = 2.07; $p = .001$) and the group with high certainty and a resource-rational chunk of length four (M = 3.23, SD = 1.69; $p = .006$).

Together, these results suggest that participants were sensitive to the resource-rational chunk length, with longer chunks being learned more readily when they were optimal, while lower transition certainty reduced the extent to which longer chunks were adopted.

Early vs. late uncertainty in action transitions Hypothesis 4 examined whether the position of certain and uncertain action transitions affects the length of chunks created by participants and predicted that if uncertainty occurred late ($0.99 \rightarrow 0.8$ or $0.99 \rightarrow 0.99 \rightarrow 0.8$; conditions 3 and 6), participants would form shorter chunks, whereas if uncertainty occurred early ($0.8 \rightarrow 0.99$ or $0.8 \rightarrow 0.99 \rightarrow 0.99$; conditions 1 and 4), participants would form longer chunks. This is because in the latter case, a longer action chunk was useful as often as a shorter action chunk, making the longer action chunk resource-rational (as it reduces planning costs more).

Because this hypothesis depended on a different grouping of a subset of the data since only conditions 1, 3, 4, and 6 are considered as opposed to the full dataset, we again examined the normality and homogeneity of the data, as preregistered. Due to the non-normality (all $W < 0.73, p < .001$) despite non-significant Levene's test for homogeneity of variances ($W = 0.18, p = .676$) of the data, we again turned to the preregistered non-parametric tests.

A two-sided Mann-Whitney U test comparing the average chunk lengths between the two transition probability timing conditions yielded a non-significant result (early uncertainty: M = 3.02, SD = 1.71; late uncertainty: M = 2.85, SD = 1.72; U = 4504.50, $p = .225$), failing to support hypothesis 4. As an exploratory analysis, we again analysed the data broken down by resource-rational action sequence length. However, these analyses did not yield significant results (see the lower right subfigure in Figure 2). In environments where a length-three action sequence was optimal, participants did not form shorter or longer abstractions in environments in which transition uncertainty occurred early (M = 2.45, SD = 0.91) than in environments in which transition uncertainty occurred late (M = 2.48, SD = 1.32; two-sided U = 1215, $p = .367$). Similar results were found for the length-four environments (early uncertainty: M = 3.67, SD = 2.13; late uncertainty: M = 3.23, SD = 2.00; two-sided U = 1078.50, $p = .251$).

Additional exploratory analyses

Frequency of resource-rational chunks To examine whether participants learned the resource-rational action sequences, we analysed the proportion of chunks in the last

Table 2: Proportion of (subsequence of) the resource-rational chunks for each condition.

Cond.	Proportion of exact resource-rational chunk	Proportion of chunks consisting of part of the resource-rational chunk
1	27.54%	64.66%
2	35.00%	76.44%
3	18.92%	83.82%
4	8.08%	61.27%
5	18.54%	67.75%
6	15.07%	74.84%

15 trials. Across all six conditions, participants used the full resource-rational chunk in 21.07% (SD: 38.05%) of the trials (see Table 2 for details on each condition and Figure 3 for the six most frequently used chunks grouped by the resource-rational chunk). The proportion of exact matches varied across conditions, ranging from 8.08% in Condition 4 (0.8 $\rightarrow$ 0.99 $\rightarrow$ 0.99) to 35% in Condition 2 (0.99 $\rightarrow$ 0.99). When considering sequences that contained parts of the resource-rational chunk, for example *jump*, *turn right* when the resource-rational chunk is *jump*, *turn right*, *jump*, the overall proportion increased substantially to 68.50% (SD: 42.67%), indicating that participants learned to use subsequences of the optimal sequence even if they did not learn to use the full chunk. In addition, we found that resource-rational chunks that consisted of two consecutive jumps (that is, *jump* $\rightarrow$ *jump* $\rightarrow$ *turn right* and *jump* $\rightarrow$ *jump* $\rightarrow$ *turn right* $\rightarrow$ *walk* as opposed to *jump* $\rightarrow$ *turn right* $\rightarrow$ *jump* and *jump* $\rightarrow$ *turn right* $\rightarrow$ *jump* $\rightarrow$ *walk*) were used in their entirety significantly more often (χ^2 = 214.78, p < .001).

These findings suggest that participants were sensitive to the structure of the resource-rational sequences. Although they did not always acquire the full resource-rational action abstraction, participants often identified subsequences, with a preference for those involving repetitive jumps.

Within-trial development While the previous analyses examined the effects of sequence length, transition probabilities, and the frequency of learned sequences, they did not consider how sequences developed within trials. Examination of within-trial dynamics revealed that most sequence adjustments involved gradual increases in length. Of all modifications, 60.34% involved adding atomic actions to the chunk (see Table 3 for an overview on all available modifications). Certain actions (pressing "Reset Memory to Previous Trial Value" or "Clear Memory") could theoretically have more of an effect than actions which affect only one instruction. To control for this in this exploratory analysis, we looked at the Levenshtein distance between modifications and found that adding atomic actions still contributed most to changes in the sequence (51.71%). This suggests that participants typically began with shorter sequences and incrementally added ac-

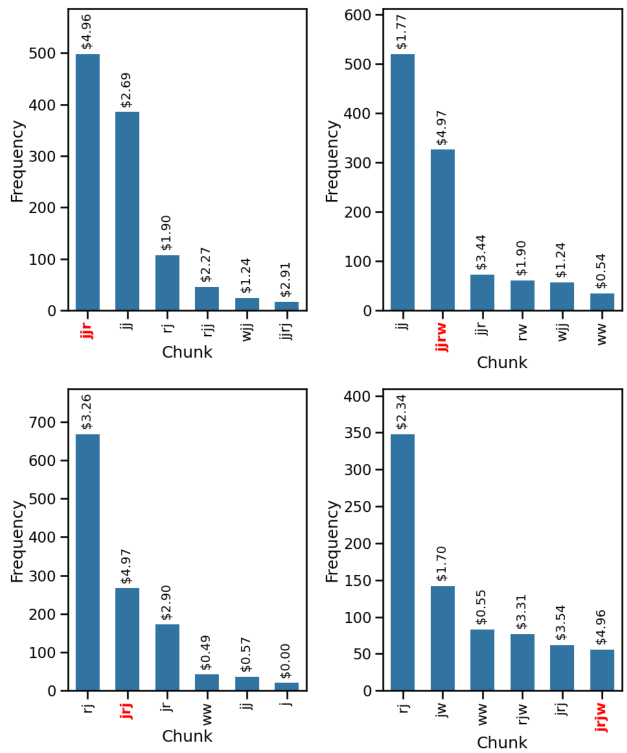

Figure 3: Aggregated frequencies of the six most frequently used chunks in the final 15 trials, grouped by map set. The chunk highlighted in red is the optimal chunk for that map set. For example, the top left subfigure shows how often the chunk *jump, jump, turn right* was used when it is the resource-rational chunk for the given map set. *J* refers to the atomic action of *jumping*, *w* refers to *walking*, *r* refers to *turn right*, and *l* refers to *turn left*. The values above each bar indicate the bonus earned by consistently using each chunk throughout the whole experiment.

tions, rather than removing actions from longer sequences or generating entirely new sequences. This pattern suggests that the acquisition of resource-rational sequences is primarily an additive process.

Discussion

This study investigated how the statistical structure of an environment shapes the formation and use of higher-level action abstractions (*chunks*) for planning. Across all six conditions, participants adapted their chunking behaviour to the environmental structure to some extent, showing systematic sensitivity to the resource-rational length of action sequences but not to differences in transition probabilities. This provides tentative evidence that people adaptively simplify planning problems in a resource-rational manner and are sensitive to environmental statistics, consistent with several previous studies (Correa et al., 2025; Eckstein & Collins, 2021; Ho et al., 2022; Solway et al., 2014). More than two thirds of the par-

Table 3: Frequency of modifications to a chunk

Type of chunk modification	Frequency	Levenshtein %
Add atomic action to *Memory*	60.34%	51.71%
Remove atomic action from *Memory*	28.73%	24.62%
Move atomic action within *Memory*	5.13%	8.80%
Press “Clear Memory”	3.18%	11.96%
Press “Reset Memory to Previous Trial Value”	2.62%	2.91%

ticipants converged on a subsequence of the resource-rational action sequence and one fifth on the exact sequence, suggesting that these subsequences still captured useful structure in the environment and yielded performance benefits, even if they did not always lead to the maximum possible bonus given the task’s reward structure.

In addition, exploratory analyses revealed systematic patterns in how action abstractions evolved over time. In particular, participants tended to modify their learned sequences by extending them, rather than by abandoning existing chunks. This incremental growth pattern suggests a learning process in which candidate abstractions are explored in depth and elaborated upon as evidence accumulates, rather than being frequently reset. One possibility is that this reflects a depth-first, forward-search-like mechanism in which an initially useful abstraction is progressively refined. Future work could test this hypothesis by developing computational models that explicitly capture incremental sequence extension and evaluating the extent to which the models’ predictions align with the observed within-trial learning dynamics.

Beyond these insights, several limitations of the present paradigm warrant consideration. While the tracing paradigm renders chunk learning explicit and observable, it also forces participants to think about chunk creation consciously, which may reduce the ecological validity of our findings. Furthermore, a central claim of chunk learning is that it reduces planning costs (Correa et al., 2025; Maisto et al., 2015). This was also made explicit in our paradigm through the remaining energy of the robot. Yet we could not measure participants’ actual planning costs. Although our design does not directly test this claim, it demonstrates that participants can discover and use structured chunks even when chunk implementation is effortful (since chunks had to be actively constructed).

Furthermore, the current paradigm does not allow us to disentangle planning-based chunking from simpler forms of statistical sequence learning or action caching. As such, our findings should be interpreted as evidence that participants learn and reuse recurring action sequences, rather than as direct evidence for resource-rational planning or reductions in cognitive planning cost. Future work can build on these findings by examining chunk learning in paradigms where chunking occurs more implicitly.

The present findings also highlight the importance of considering heterogeneity in the cognitive costs of individual atomic actions. While some action combinations may be combined with little effort, like opening a door, others may impose greater cognitive demands. In the context of the Lightbot task, actions such as turning left or right may be more cognitively demanding because they require spatial transformations and mental rotation, processes that are known to be challenging for many individuals. Assigning different cognitive costs to different atomic actions may therefore provide a more realistic account of how people form higher-level abstractions, and future studies could directly manipulate or model such asymmetries.

Relatedly, we observed that certain action sequences were more readily incorporated into learned abstractions than others. In particular, sequences involving repeated actions (e.g., consecutive jumps) were learned more often, independent of transition probabilities. One possibility is that reusing the same action within a sequence reduces retrieval costs, making such sequences easier to encode and deploy as a single unit. This suggests that chunk formation may be constrained not only by environmental statistics but also by regularities within the action sequence itself. Future work could investigate this hypothesis by systematically varying action repetition and measuring its impact on abstraction learning.

Additionally, some of our descriptive metrics—such as the proportion of participants whose behaviour matched or contained subsequences of the resource-rational chunk—should be interpreted with caution, as we did not establish an appropriate chance baseline given the limited action space. This makes it difficult to assess the extent to which these patterns exceed what would be expected under simple action reuse or random composition.

Overall, this work demonstrates that human action abstraction during planning is likely to be shaped by the statistical structure of the task. By showing how people adapt their action abstraction in response to these factors, this study advances our understanding of planning as a resource-rational process.

Acknowledgments

We acknowledge the use of large language model (LLM) tools for supporting smaller coding tasks and assisting in brainstorming appropriate statistical tests and editing. All final decisions, interpretations, and code validation were conducted by the authors. We thank Martin Butz for helpful comments and discussions, and Hanqi Zhou for additional support. We are grateful to Jean-Claude Passy for assistance in setting up the experiment and for valuable administrative support throughout the project. This work was funded by the Max Planck Institute for Intelligent Systems Grassroots Program. The authors also thank the International Max Planck Research School for Intelligent Systems (IMPRS-IS) for supporting Maximilian Mittenbühler.

References

Badre, D., Kayser, A. S., & D'Esposito, M. (2010). Frontal Cortex and the Discovery of Abstract Action Rules. *Neuron*, *66*(2), 315–326. https://doi.org/10.1016/j.neuron.2010.03.025

Botvinick, M. M., Niv, Y., & Barto, A. G. (2009). Hierarchically organized behavior and its neural foundations: A reinforcement learning perspective. *Cognition*, *113*(3), 262–280.

Correa, C. G., Sanborn, S., Ho, M. K., Callaway, F., Daw, N. D., & Griffiths, T. L. (2025). Exploring the hierarchical structure of human plans via program generation. *Cognition*, *255*, 105990.

Dezfouli, A., & Balleine, B. W. (2013). Actions, action sequences and habits: Evidence that goal-directed and habitual action control are hierarchically organized. *PLoS Computational Biology*, *9*(12), e1003364.

Eckstein, M. K., & Collins, A. G. (2021). How the mind creates structure: Hierarchical learning of action sequences. *Proceedings of the Annual Meeting of the Cognitive Science Society*, *43*, 618.

Éltető, N., & Dayan, P. (2023). Habits of mind: Reusing action sequences for efficient planning. *Proceedings of the Annual Meeting of the Cognitive Science Society*, *45*(45).

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.

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*, e1.

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.

Niv, Y. (2019). Learning task-state representations. *Nature Neuroscience*, *22*(10), 1544–1553.

Palan, S., & Schitter, C. (2018). Prolific.ac—a subject pool for online experiments. *Journal of Behavioral and Experimental Finance*, *17*, 22–27.

Sanborn, S., Bourgin, D. D., Chang, M., & Griffiths, T. L. (2018). Representational efficiency outweighs action efficiency in human program induction. *Proceedings of the Annual Meeting of the Cognitive Science Society*, *40*.

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.

Sutton, R. S., Precup, D., & Singh, S. (1999). Between mdps and semi-mdps: A framework for temporal abstraction in reinforcement learning. *Artificial Intelligence*, *112*(1-2), 181–211.

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.